THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

Systèmes Informatiques

Présentée et soutenue par :

Mme MAIALEN LARRANAGA

le vendredi 25 septembre 2015

Titre :

DYNAMIC CONTROL OF STOCHASTIC AND FLUID RESOURCE-SHARING SYSTEMS

Ecole doctorale :

Systèmes (Systèmes)

Unité de recherche :

Laboratoire d'Analyse et d'Architecture des Systèmes (L.A.A.S.)

Directeur(s) de Thèse :

M. URTZI AYESTA

M. FRANTZISKO XABIER ALBIZURI IRIGOYEN

Rapporteurs :

M. ALAIN JEAN-MARIE, UNIVERSITE MONTPELLIER 2

M. GER KOOLE, VRIJE UNIVERSITEIT AMSTERDAM

Membre(s) du jury :

M. ANDRE LUC BEYLOT, INP TOULOUSE, Président

M. FRANTZISKO XABIER ALBIZURI IRIGOYEN, UNIVERSIDAD DEL PAIS VASCO, Membre

M. KEVIN GLAZEBROOK, UNIVERSITE DE LANCASTER, Membre

Mme INA MARIA VERLOOP, INP TOULOUSE, Membre

M. RUDESINDO NUNEZ-QUEIJA, CENTRUM WISKUNDE INFORMATICA AMSTERDAM, Membre

M. URTZI AYESTA, LAAS TOULOUSE, Membre

Acknowledgements

Words cannot express the gratitude I feel towards my advisors Urtzi and Maaike. I have enjoyed every phase of this three years long process. I have learned a lot from both of them. I thank them for the environment that I have been offered built on respect, trust and generosity. I also thank Xabier for establishing the collaboration with the University of the Basque Country and for being this helpful all along the process.

I thank Airbus Group Corporate Foundation for funding my Ph.D. thesis and the conference trips.

During a two months research visit I was able to work in CWI and Eurandom with Rudesindo Nuñez-Queija and Onno Boxma. My sincere thanks to both of them for hosting me and including me in their corresponding research groups. I enjoyed the short but intense visits very much. These research visits were funded by Edsys and INP, for which I thank both establishments.

I thank the jury members André-Luc Beylot, Kevin Glazebrook, Ger Koole, and Rudesindo Nuñez-Queija, for having given me feedback on my thesis and for the enriching discussions that took place the day of the defense. Also I thank Alain Jean-Marie, who unfortunately could not make it to the defense, but has thoroughly read the thesis and the given feedback has helped improve the final version.

I thank my fellow labmates Ahmad, Tania, Henda, Christo, Josselin, Josu, Bala and Olivier for the fun times during lunch and coffee breaks. I shared these three years journey with many people that have made a difference, my office mates in LAAS Tom and Samir, it has been a real pleasure to be part of E39, I enjoyed all the craziness and fun in that office. I also thank my office mates in ENSEEIHT Farouk and Chakadkit. I thank Ane and Martin, my scientific sister and brother and I wish them success in their future projects. Even though I have not seen them much during the past three years I thank Imanol, Tijana, Maria, Julia, Felipe, Javier and Vincent, with whom I shared my first steps into research.

I would like to warmly thank Sylvie, Sonia and Caroline for dealing with late conference registrations and all the trip arrangements.

Quiero agradecer también a los amigos con los que he compartido penas y alegrías en el LAAS y alrededores: Idoia, Aloña, Guillaume, Isidre, Carlos (bis), Nuria, Oswaldo, Paqui, Grego y Marti, gracias.

Je remercie Denis de ses encouragements, de m'avoir supportée pendant les moments plus stressants et surtout de son soutien quotidien sans faille. Un grand merci.

Eskerrak Amaia eta Maiteri, sopresa zoragarri bat eman izanagatik eta zeraten bezain alaiak izateagatik. Azkenik guraso, ahizpa eta familia eskertu nahi ditut. Azken urte hauetan gutxi xamar ikusi nauzuen arren beti hor izan zaretelako, eskerrak bihotz-bihotzez.

Abstract

In this thesis we study the dynamic control of resource-sharing systems that arise in various domains: *e.g.* inventory management, healthcare and communication networks. We aim at efficiently allocating the available resources among competing projects according to a certain performance criteria. These type of problems have a stochastic nature and may be very complex to solve. We therefore focus on developing well-performing heuristics. In Part I, we consider the framework of Restless Bandit Problems, which is a general class of dynamic stochastic optimization problems. Relaxing the sample-path constraint in the optimization problem enables to define an index-based heuristic for the original constrained model, the so-called Whittle index policy. We derive a closed-form expression for the Whittle index as a function of the steady-state probabilities for the case in which *bandits* (projects) evolve in a birth-and-death fashion. This expression requires several technical conditions to be verified, and in addition, it can only be computed explicitly in specific cases. In the particular case of a multi-class abandonment queue, we further prove that the Whittle index policy is asymptotically optimal in the light-traffic and heavy-traffic regimes. In Part II, we derive heuristics by approximating the stochastic resource-sharing systems with deterministic fluid models. We first formulate a fluid version of the relaxed optimization problem introduced in Part I, and we develop a fluid index policy. The fluid index can always be computed explicitly and hence overcomes the technical issues that arise when calculating the Whittle index. We apply the Whittle index and the fluid index policies to several systems: *e.g.* power-aware server-farms, opportunistic scheduling in wireless systems, and make-to-stock problems with perishable items. We show numerically that both index policies are nearly optimal. Secondly, we study the optimal scheduling control for the fluid version of a multi-class abandonment queue. We derive the fluid optimal control when there are two classes of customers competing for a single resource. Based on the insights provided by this result we build a heuristic for the general multi-class setting. This heuristic shows near-optimal performance when applied to the original stochastic model for high workloads. In Part III, we further investigate the abandonment phenomena in the context of a content delivery problem. We characterize an optimal grouping policy so that requests, which are impatient, are efficiently transmitted in a multi-cast mode.

Contents

1 Introduction **1**
1.1 Brief background on queueing theory . . . 2
1.2 Modeling framework . . . 2
1.2.1 Restless bandit problem . . . 3
1.2.2 Examples of multi-class allocation problems . . . 4
1.2.3 Abandonment phenomena . . . 6
1.3 Methodology . . . 8
1.3.1 Lagrangian relaxation . . . 8
1.3.2 Fluid approximation . . . 10
1.3.3 Optimal stochastic control . . . 10
1.3.4 Optimal deterministic control . . . 13
1.4 Overview of the thesis . . . 15

I Dynamic control of stochastic resource-sharing systems **17**

2 Index policy for birth-and-death restless bandits **19**
2.1 Introduction . . . 19
2.2 Model description . . . 20
2.2.1 Examples . . . 21
2.3 Relaxation and indexability . . . 22
2.3.1 Threshold policies . . . 23
2.3.2 Indexability and Whittle's index . . . 24
2.4 Whittle's index policy . . . 27
2.5 Performance of Whittle's index policy . . . 27
2.6 Appendix . . . 28
2.6.1 Proof of Proposition 2.1 . . . 28
2.6.2 Proof of Proposition 2.2 . . . 28
2.6.3 Proof of Corollary 2.1 . . . 30

3 Index policy for an abandonment queue **33**
3.1 Model description and preliminaries . . . 34
3.2 Relaxation . . . 36
3.2.1 Threshold policies . . . 36
3.2.2 Indexability . . . 38
3.2.3 Whittle's index . . . 38
3.3 Case studies . . . 39
3.3.1 Linear holding cost . . . 39
3.3.2 Convex holding cost . . . 41
3.4 Multi-class M/M/1 queue . . . 43
3.5 Asymptotic optimality . . . 45
3.5.1 Multi-server setting . . . 45
3.5.2 Light-traffic and heavy-traffic regimes . . . 45
3.6 Numerical results . . . 46
3.6.1 Structure of different policies . . . 46
3.6.2 Performance evaluation . . . 47
3.7 Appendix . . . 51
3.7.1 Proof of Proposition 3.1 . . . 51
3.7.2 Proof of Proposition 3.3 . . . 55
3.7.3 Proof of Proposition 3.4 . . . 56
3.7.4 Proof of Propostion 3.5 . . . 60
3.7.5 Proof of Proposition 3.6 . . . 62
3.7.6 Proof of Proposition 3.7 . . . 64
3.7.7 Proof of Proposition 3.8 . . . 66
3.7.8 Proof of Proposition 3.9 . . . 67

II Dynamic control of fluid resource-sharing systems **69**

4 Fluid index policy for birth-and-death restless bandits **71**
4.1 Fluid version of relaxed optimization problem . . . 72
4.1.1 Fluid model and bias optimality . . . 72
4.1.2 Optimal fluid control and fluid index . . . 74
4.1.3 Fluid index policy . . . 76
4.2 Equivalence of stochastic index and fluid index in light-traffic regime . . . 76
4.3 Case studies . . . 78
4.3.1 Scheduling in a multi-class abandonment queue . . . 78
4.3.2 Opportunistic scheduling in a wireless downlink channel . . . 80
4.3.3 Routing/blocking in a power-aware server-farm . . . 83
4.3.4 Scheduling a make-to-stock queue with perishable items . . . 86
4.4 Appendix . . . 90
4.4.1 Proof of Proposition 4.1 . . . 90
4.4.2 Lemma 4.2 . . . 95

4.4.3 Proof of Proposition 4.2 . . . 98
4.4.4 Proof of Proposition 4.3 . . . 102
4.4.5 Proof of Proposition 4.4 . . . 103

5 Optimal fluid control for an abandonment queue **105**
5.1 Introduction . . . 105
5.2 Model description . . . 107
5.3 Fluid control model . . . 108
5.3.1 Optimal policy in underload for two classes of customers . . . 109
5.3.2 Optimal policy in overload for an arbitrary number of classes . . . 111
5.3.3 Optimal control comparison of stochastic model with fluid model . . . 112
5.4 Heuristics for an arbitrary number of classes . . . 114
5.5 Numerical results . . . 116
5.5.1 Performance analysis for two classes of customers . . . 116
5.5.2 Performance analysis for arbitrary number of customers . . . 119
5.6 Appendix . . . 120
5.6.1 Proof of Lemma 5.1 . . . 120
5.6.2 Proof of Proposition 5.1 . . . 123

III Optimal dynamic control of stochastic systems **129**

6 Batch queues with abandonments **131**
6.1 Introduction . . . 131
6.2 Model description . . . 132
6.3 Optimality of threshold policies . . . 135
6.4 Steady-state distribution . . . 135
6.4.1 Infinite service rate . . . 136
6.4.2 Finite service rate . . . 136
6.5 Characterization of the optimal threshold . . . 138
6.6 Examples . . . 139
6.7 Appendix . . . 142
6.7.1 Proof of Proposition 6.1: the case $\mu < \infty$ . . . 142
6.7.2 Proof of Proposition 6.2 . . . 146
6.7.3 Proof of Proposition 6.3 . . . 150
6.7.4 Proof of Lemma 6.1 . . . 151
6.7.5 Justification of Conjecture 6.1 . . . 153

A Appendix **155**
A.1 Conditions for the existence of optimal stationary policies: The Bellman equation . . . 155
A.2 Necessary conditions: Pontryagin's Minimum Principle . . . 156
A.3 Sufficient conditions: Hamilton-Jacobi-Bellman Equation . . . 157

Résumé en français **161**

Laburpena euskaraz **165**

Self References **167**

Bibliography **169**

Chapter 1

Introduction

Contents

1.1 **Brief background on queueing theory** 2
1.2 **Modeling framework** 2
1.3 **Methodology** 8
1.4 **Overview of the thesis** 15

The problem of allocating resources to heterogeneous competing classes of customers has received a lot of attention from the research community due to its applications into various domains, *e.g.*, inventory management, healthcare, communication networks. Resource allocation policies may be designed to achieve different objectives: *e.g.*, maximization of sales in stock management, minimization of casualties in triage systems or maximization of throughput in wireless networks. From a mathematical point of view, queueing-theoretic models are able to capture many of the features of resource-sharing systems, and they have traditionally been considered to evaluate the performance of queueing systems, *e.g.*, predict the queue-length distribution of customers, compute the sojourn times, as well as to help in the decision making process, *e.g.*, develop scheduling policies to achieve the desired goal.

Many of the resource-sharing systems have a random nature and therefore in this thesis we have focused on stochastic control models, more specifically, on Continuous Time Markov Decision Processes (CTMDPs). In some instances it is possible to determine the optimal control. However, in most cases finding an optimal solution is out of reach and one is bound to seek well-performing heuristics. In this thesis we study a wide range of resource-sharing systems, we characterize an optimal solution whenever possible and we develop heuristics when the optimal solution cannot be found.

In this chapter we present a brief background on queueing theory, see Section 1.1, we motivate the modeling framework considered throughout the thesis, see Section 1.2, we introduce the methodology that will be used in the thesis, see Section 1.3 and finally, we present an overview of the thesis, see Section 1.4.

1.1 Brief background on queueing theory

Queueing systems were initially developed in the area of telecommunications. A. K. Erlang in [43], a Danish mathematician, used them to represent the telephone exchange. Queueing models have been widely used to analyze many complex problems arising in Operations Research. The origin of modern Operations Research dates back to World War II, when the British government used this discipline to solve operational military problems. Later it has spread to transportation, manufacturing industry and healthcare among others. In the late 50s, when the first computers showed up, queueing models started being used for evaluating the performance of computer systems and solve complex decision making problems. Over the years Stochastic Operations Research has become extremely popular among computer scientists and applied mathematicians to solve problems that emerge from wireless networks, social networks, data transfers, traffic management, etc. As an illustration of the increasing influence of queueing theory in todays research we have the journal Queueing Systems established in 1986 which is entirely devoted to it. Queueing-theoretic models, although they have been used to model realistic systems, are rather *simple* and yield to mathematical analysis. This leads to an understanding on the performance of different systems. For an overview on queueing theory we refer to Kleinrock [63, 64] and Takagi [88].

A queueing model is usually specified by the following elements: an input process, a departure process, and a service discipline. The input process establishes the arrivals of customers to the system, the departure process defines the time it takes to serve a customer, and the discipline decides how the customers in the queue are chosen to be served. These processes may have a random nature. Kendall, in [61], introduced the following notation to describe queueing systems

$$A/B/s/C,$$

where A denotes the distribution of the inter-arrival times, B the distribution of the service times, s the amount of servers that are available in the system and C the total capacity of the system, *i.e.*, maximum number of customers that are allowed in the system. A semantic remark is in order, namely, the term *customers* will be used throughout the thesis to refer to demands for a specific item, tasks, patients or users, and the term *servers* to refer to entities with processing capacity, for instance, doctors, operators or machines.

Queues are typically modeled by Markov chains. Markov chains are sequences of random events that only depend on the current state of the system, this property is better known as *memoryless*. Markov chains have been proved to be a good approximation to various systems in physics, in biology to model population processes, speech recognition, etc. The canonical example of Markovian queues is the M/M/1 queue, where M refers to exponential distribution of inter-arrival times and service times, and 1 for a single server. This model has received a lot of attention in the literature and will play an important role in this thesis, we refer to the monograph Cohen [36] for an overview on single-server queues.

1.2 Modeling framework

In this thesis we study the dynamic control of resource-sharing systems. We consider the *restless bandit problems* (RBPs) framework, which is a general class of resource-sharing problems, see Section 1.2.1. In particular, we focus on two multi-class allocation problems that fall within the RBP setting, that of optimal class selection and that of load balancing, see Section 1.2.2. In the thesis a special interest is given to resource-sharing problems in the presence of customers abandonments. In Section 1.2.3, we therefore provide a discussion on the abandonment phenomena and the particular models in which we have considered it.

1.2.1 Restless bandit problem

The framework of RBPs is a special class of Markov Decision Processes (MDPs) or stochastic dynamic optimization. The basic model can be described as follows. There are K bandits in the system, and at most M bandits can be activated. The objective is to find a policy that optimally activates the bandits in order to obtain a certain performance criteria. Bandit k is then specified by the following elements: a state space $E_k = \{0, 1, \ldots\}$, an action space $A_k = \{0, 1\}$, where 0 represents that bandit k is *passive* and 1 represents that bandit k is *active*, the transition rates $q_k^a(m, \tilde{m})$ with $m, \tilde{m} \in E_k$ and $a \in \mathcal{A}_k$, and a cost function $C_k(m, a)$ for all $k = 0, 1, \ldots K$, which depends both on the state of the system $m \in E_k$ and the action $a \in \mathcal{A}_k$. Bandits depend on each other through the control policy since at most M bandits can be made active at a time.

We let the state of the system be defined by $\vec{N}^\phi(t) := (N_1^\phi(t), \ldots, N_K^\phi(t))$, where $N_k^\phi(t) \in E_k$ represents the state of bandit k under policy ϕ. Let us further denote by $S_k^\phi(\vec{N}^\phi(t)) \in \{0, 1\}$, the action taken with respect to bandit k when the state of the system is $\vec{N}^\phi(t)$. The objective is to obtain the allocation policy ϕ that minimizes the long-run expected average cost incurred by the system

$$\limsup_{T \to \infty} \frac{1}{T} \mathbb{E} \left(\int_0^T \sum_{k=1}^K C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t)))\mathrm{d}t \right). \tag{1.2.1}$$

In addition, the following constraint on available resources must be satisfied at all decision epochs:

$$\sum_{k=1}^K S_k^\phi(\vec{N}^\phi(t)) \leq M. \tag{1.2.2}$$

The RBP framework was first introduced by Whittle in [99] as a generalization of the Multi-Armed Bandit Problem (MABP). In a MABP problem it is assumed that $M = 1$, that is, at every decision epoch the scheduler needs to activate one bandit. The activated bandit incurs a cost and its state evolves stochastically, while the states of all other bandits remain *frozen*. In a ground-breaking result Gittins showed that the optimal policy that solves a MABP is an index rule, nowadays commonly referred to as Gittins' index policy [48, Chapter 2.5]. An index rule or policy is such that there exist functions $G_k(m_k)$ for all $k \in \{1, \ldots, K\}$, depending only on the parameters of bandit k, for which the optimal policy in state $\vec{m}$ prescribes to serve the bandit having currently the highest index $G_k(m_k)$. In multiple problems of practical interest, the problem cannot be cast as a MABP, that is, bandits also evolve and incur a cost even though they have not been activated by the scheduler, this is modeled by the RBP setting. We note that in the RBP context, in general, an optimal policy cannot be characterized.

The RBP theory has gained a lot of popularity thanks to the important contributions of researchers like K. Glazebrook, J. Niño-Mora and R. Weber, see Glazebrook *et al.* [5, 48, 50], Niño-Mora *et al.* [23, 71], Weber *et al.*[96] for a few examples.

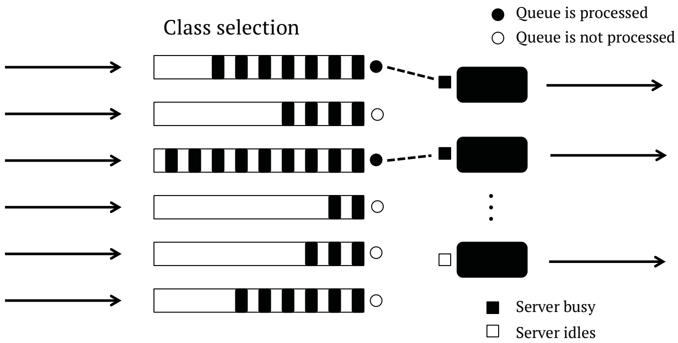


Figure 1.1: The problem of allocating resources to classes of customers

In this thesis we are going to focus on RBP problems in which bandits follow a very specific dynamic. The first type of processes that we are going to consider are birth-and-death processes, namely, each bandit has two possible transitions: *births*, that is, $q_k^a(m, m+1) > 0$ for $m \in E_k$, or *deaths*, that is, $q_k^a(m, m-1) > 0$ for $m \in \{1, 2, \ldots\}$. The second type of processes that we consider are models with transitions that can either be *births*, *deaths*, or *batch departures* of length m, that is, $q_k^a(m, 0) > 0$, with $m \in E_k$. The latter will be considered in Chapter 6, where we study a server with infinite capacity and upon activation of service all customers waiting in the queue receive service at the same time.

The systems that we analyze throughout this thesis are assumed to be ergodic, that is, we focus on the set of stationary policies under which a limiting stationary distribution can be found for the underlying Markov processes. For an overview on ergodicity of Markov processes we refer to Guo *et al.* [54, Appendices B and C].

1.2.2 Examples of multi-class allocation problems

In this section we first describe two classes of examples that fall within the RBP framework decribed in Section 1.2.1, and that are of particular interest in this thesis. Namely, that of optimal class selection and that of load balancing. Second, we present a brief discussion of the cost functions that we consider in this thesis.

Optimal class selection problems

In optimal class selection problems, customers of a given class k arrive to their corresponding queue. In this setting the resources that are available are servers itself, which we consider to be M. The objective is to allocate these M servers to the K competing classes of customers. Each server can be dedicated to at most one class, and each class can receive service from at most one server. A class-k customer departs after an exponentially distributed time, where the transition rate depends on whether a server is allocated to class k and on the state of bandit k (*i.e.,* number of customers). The decision on how to allocate the servers will be based on the holding cost of each class and other extra model dependent costs. For an illustration of optimal class selection problems see Figure 1.1. This is a special class of RBPs: a bandit represents a class of customers and the state of bandit k is the number of class-k customers. In particular, the evolution of a bandit is of birth-and-death type. In this setting the action variable $S_k^{\phi}(\vec{N}^{\phi}(t))$ describes

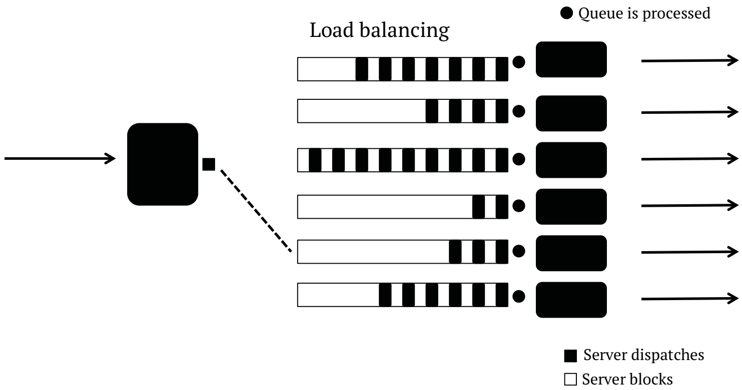


Figure 1.2: The problem of balancing the load in heterogeneous queues

whether bandit k is made active or not, *i.e.*, $S_k^\phi(\vec{N}^\phi(t)) = 1$ if class k is served and $S_k^\phi(\vec{N}^\phi(t)) = 0$ if class k is not served. In Section 2.2.1 this model will be explained in further detail.

This framework is used for instance to model call centers. In call centers customers of a given class, where a class might represent the type of information or help they require, are held until an operator is ready to process their requests. Due to the limited amount of operators, customers of different classes might need to wait. The call center needs to make a choice on which class of customers must be prioritized in order to maximize the revenue or the customers experience.

In this setting we will consider three type of problems that fit the framework of optimal class selection problems: a multi-class single-server abandonment queue (see Chapter 3, Section 4.3.1 and Chapter 5), a wireless downlink channel under opportunistic scheduling (see Section 4.3.2) and a multi-cast transmission queue in the presence of customers abandonment (see Chapter 6).

Load balancing problems

In load balancing problems, customers that arrive to the system are being dispatched to K different queues. The dispatchers decide to which of the K queues to route the newly arrived customer, or to block the customer. In this case the dispatchers play the role of the resource. Throughout this thesis we will consider that in load balancing problems a single dispatcher is available, that is, $M = 1$. Furthermore, once a customer has been dispatched to one of the K queues it departs after an exponentially distributed amount of time with a rate that depends on the number of customers present in that queue. The decision to be made is then whether an arriving customer is accepted to the system and routed to one of the queues. This problem falls also within the RBP context, where a bandit represents the number of customers in a queue, and the state of bandit k is the number of customers in queue k. The evolution of a bandit is of birth-and-death type. An arriving customer is routed to queue k (*i.e.,* bandit k is activated) if there exists $k \in \{1, \ldots, K\}$ such that $S_k^\phi(\vec{N}^\phi(t)) = 1$ and $S_j^\phi(\vec{N}^\phi(t)) = 0$ for all $j \neq k$, or it is blocked if $S_k^\phi(\vec{N}^\phi(t)) = 0$ for all $k \in \{1, \ldots, K\}$ (*i.e.,* no bandit is activated). This mechanism allows to balance the workload in the servers, see Figure 1.2 for an illustration. In Section 2.2.1 this model will be explained in further detail.

A classical example in load balancing problems is routing in server farms. K different servers are active to process tasks, tasks arrive to the system where the dispatcher needs to route them to one of the servers. If the task has been routed to a busy server then it lines up until the server becomes available for processing.

Contrary to the optimal class selection problems, in load balancing problems the decision is made at the dispatcher level and not at the server level. In this thesis we will analyze two different models that fall in the load balancing framework: that of scheduling in a power-aware server farm, see Section 4.3.3, and inventory management in a make-to-stock problem with perishable items, see Section 4.3.4.

Assumptions on the cost function

Having introduced the general modeling framework in Section 1.2.1, we make assumptions on the cost function. The cost function $C_k(m, a)$ for all $k \in \{1, \dots, K\}$, $m \in E_k$ and $a \in \{0, 1\}$ will be model dependent, but in general it will consist of the sum of the holding cost and an extra cost coming from the particular phenomena considered in the models. The latter can be: (i) cost per customer abandonment, when a customer decides to leave the system before service completion, (ii) cost for power consumption, the more customers are held in the queue the faster the server processes customers, which is a consequence of the speed-scaling rule, (iii) a cost per blocked customers, in some systems it is preferred not to let customers enter the system rather than paying a cost for holding them in the queue, and (iv) set-up cost, when setting the service up a cost is incurred. The latter four phenomena will be studied in the thesis. We will assume that $C_k(m, a)$ is convex in m for $a = 0, 1$, for all $k \in \{1, \dots, K\}$. Convex holding cost have been considered to model many queueing systems and networks, see Ansell *et al.* [5], Dai *et al.* [39, 40] and Weber *et al.* [95] for a few examples. In [70] Van Mieghem argued that linear delay holding cost were not suitable in many settings. By considering convex holding cost one can for example capture the credibility of customers or the market reputation. In [4] Ansell *et al.* highlight the disadvantages of minimizing linear holding cost in a two-class single-server queue, since the lowest priority queue grows large.

1.2.3 Abandonment phenomena

In this thesis a special emphasis is put on models considering the phenomena of customers abandonment. Abandonment or reneging takes place when customers, unsatisfied of their long waiting time, decide to voluntarily leave the system. It has a huge impact in various real life applications such as the Internet or call centers, where customers may abandon while waiting in the queue, or even while being served. Abandonment is a very undesirable phenomena, both from the customers' and system's point of view (a profit is lost), and it can have a big economical impact. It is thus not surprising that it has attracted considerable interest from the research community, with a surge in recent years. To illustrate this, we can mention the recent Special Issue in Queueing Systems on queueing systems with abandonments [57] and the survey paper by Dai *et al.* [38] on abandonments in a many-server setting. More particularly, an important line of research aims at characterizing the performance and the impact of abandonments in systems, we refer to Ata *et al.* [7], Baccelli *et al.* [16], Brandt *et al.* [30], Brill *et al.* [31], Gromoll *et al.* [53] and Hasenbein and Perry [58] for single-server models and Boots *et al.* [27], Boxma *et al.* [29] and Whitt [98] for papers dealing with the multi-class case. Related literature that is more close to the work in this thesis consists of papers that deal with optimal scheduling or control aspects, see for instance Argon *et al.* [6], Ata *et al.* [7], Atar *et al.* [8, 9], Ayesta *et al.* [15], Bhulai *et al.* [24], Down *et al.* [41], Glazebrook *et al.* [51] and Kim *et al.* [62].

We now briefly introduce the particular models in the presence of customers impatience that have been analyzed in this thesis.

Multi-class single-server abandonment queue

In this model we assume that K classes of customers are competing for one single resource, and customers that are either waiting in the queue or are being served may abandon the system. This model belongs to the class of problems depicted in Figure 1.1, with $M = 1$. For each class, the arrival process is assumed to be a Poisson process and the service requirements and the abandonment times are exponentially distributed. The objective is to design policies that minimize the average cost in steady-state. A cost is incurred every unit of time customers are held in the queue and per customer abandonment. This multi-class single-server abandonment model has for instance been considered in Bhulai *et al.* [24] and Salch *et al.* [81, 82]. When deciding which class to serve, a scheduler chooses between myopically minimizing the cost and not serving the class with high abandonment rate. The system might profit from the latter by letting the classes of customers with high abandonment rates grow large (without idling).

In general, determining the exact optimal policy has so far proved analytically infeasible. Hence, we have focused on obtaining approximations for the optimal control. In this thesis we consider two different approaches to approximate this model, the first approach is based on the Lagrangian relaxation method, which is performed in Chapter 3, and the second approach is based on a fluid approximation, which is performed in Chapter 5. The general theory of both approximation methods can be found in Section 1.3.1 and Section 1.3.2, respectively.

Make-to-stock with perishable items

In this model we consider an inventory management company which produces K different classes of items and stores them until they are sold or they perish. For simplicity we consider $M = 1$, that is, a single machine produces all the items. When a demand for an item k arrives to the company there will be a sale if there are items of class k in stock. However, if no item of class k is stocked, then the sale will be lost, incurring an extra cost. We also assume that items are perishable which can be modeled with abandonments. We aim at obtaining the stocking policy such that we minimize the holding cost together with the cost for losing a sale and cost per perished item. Due to items being perishable, it might be better to let the machine idle than produce items in excess. Characterizing an optimal solution for this problem is very complicated and we will therefore focus on obtaining well-performing heuristics. This problem belongs to the class of problems depicted in Figure 1.2.

This make-to-stock problem has previously been considered in Veatch *et al.* [91, 92] with non-perishable items. As far as we are aware, the multi-class make-to-stock problem has not been analyzed under the perishability assumption. We will analyze this model in Section 4.3.4.

Multi-cast transmission in the presence of abandonments

In large-scale, high-volume content-delivery and file-sharing networks, bandwidth resources are scarce and must be efficiently used. As the bulk of traffic is delay tolerant (*e.g.*, software updates, video content), customers that request a content can be grouped, so that the content is transmitted in multi-cast mode through the network. After a request has been received, it may be better to postpone the actual transmission until one or more additional requests for the same content arrive, which may save tremendous transmission capacity. The objective is to determine the clearing policy, that is, to decide when is the right time to take a batch into service in order to minimize the average holding and set-up cost. The trade-off in this model comes from the set-up cost, that is, establishing a multi-cast transmission might be very expensive. The challenge is to balance these costs against the risk of not meeting the deadline of one or more requests. We assume the deadlines to be exponentially distributed.

This problem can be modeled with a single bandit, where the state of the bandit represents the number of requests. The transitions rates are $q^1(m,0) > 0$ for all $m > 0$ if the bandit is active and $q^0(m,0) = 0$ for all $m > 1$ if bandit is passive, the transition $q^a(m,m+1)$ for all $m \geq 0$ corresponds to the arrival of a new request. Note that the model is a special case of the problems described in Section 1.2.2. This model yields to exact analysis, in such a way that it is possible to determine the optimal policy. We will study it in Chapter 6 where we characterize the optimal clearing policy.

1.3 Methodology

As we have discussed in the previous section, the resource-sharing models that we considered in this thesis are such that in general an optimal allocation policy can not be found or are only attainable for specific examples. Numerically, one can try to use methods for solving MDPs, such as, *value iteration* or *policy iteration*, see Bertsekas [22, Vol II, Section 4.3] and Puterman [77], however, when we perform numerical experiments we encounter several issues: unbounded transition rates, curse of dimensionality, infinite state-space, etc.

In this section we will give an overview on possible approximations to solve MDPs, which have been applied in this thesis. The objective is to obtain heuristics that are efficient for the RBP as introduced in Section 1.2.1.

In Section 1.3.1 we introduce the Lagrangian relaxation approach which allows us to simplify the multi-armed RBPs by solving simpler single-armed RBPs. In Section 1.3.2 we introduce a fluid approximation which enables the stochastic RBP to be approximated by a deterministic control problem. Once these approximations have been presented, we discuss in Sections 1.3.3 and 1.3.4 tools for deriving optimal solutions for both stochastic and deterministic control problems. For completeness, the most relevant theorems can be found in Appendix A, at the end of the thesis.

1.3.1 Lagrangian relaxation

In this section we introduce a method for deriving heuristics for RBPs. Recall the RBP introduced in Section 1.2.1, which is specified by the objective function (1.2.1) and the constraint on the available resources as in (1.2.2). In order to obtain a well-performing heuristic, in the seminal work [99] Whittle proposed to relax the constraint (1.2.2) by letting it be satisfied on average instead of at all units of time. That is, (1.2.2) is relaxed to obtain

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \sum_{k=1}^{K} S_k^{\phi}(\vec{N}^{\phi}(t))\right) \leq M. \tag{1.3.1}$$

The objective function (1.2.1) together with the relaxed constraint (1.3.1), constitute what we will denote by the *relaxed optimization problem*. Equivalently, we can write this as: minimize

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\bigg(\int_0^T \bigg(\sum_{k=1}^K C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t))) + W(\sum_{k=1}^K S_k^\phi(\vec{N}^\phi(t)) - M) \bigg) \mathrm{d}t \bigg). \tag{1.3.2}$$

This is a commonly used method to solve constrained optimization problems, and is known as the *Lagrange multiplier* technique, see Bertsekas [20], [21, Chapters 3 and 4], where W is the Lagrange multiplier. By integrating the constraint of the system in the objective function together with a multiplier, one obtains an unconstrained optimization problem, which is simpler to solve than the original RBP. We will refer to (1.3.2) as the *unconstrained optimization problem*. The Lagrange multiplier W penalizes policies for which the constraint (1.3.1) is not satisfied on average.

The relaxation of the constraint (1.2.2) together with the use of the Lagrangian multipliers method to obtain (1.3.2), constitute the *Lagrangian relaxation* technique. The key observation made by Whittle is that problem (1.3.2) can be decomposed into K subproblems, one for each different bandit k, that is, one can independently analyze the K unidimensional problems

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\bigg(\int_0^T \bigg(C_k(N_k^\phi(t), S_k^\phi(N_k^\phi(t))) + W S_k^\phi(N_k^\phi(t)) \bigg) \mathrm{d}t \bigg), \tag{1.3.3}$$

for all $k \in \{1, \ldots, K\}$. In the RBP context W can be interpreted as a penalty for activity (equivalently, as a subsidy for passivity). We define by $W_k(m_k)$ the value of W such that the optimal policy that solves Problem (1.3.3) is indifferent of the action taken in state m_k, referred to as *Whittle's index*. Whittle showed that the solution to the unconstrained optimization problem (1.3.2), provided that a so-called *indexability* property holds, is of index type with index $W_k(m_k)$. That is, it is optimal to activate all bandits for which $W_k(m_k) \geq W$, see Whittle [99]. Whittle then defined a heuristic for the original constraint problem (1.2.1)-(1.2.2), referred to as Whittle's index policy, where in every decision epoch the M bandits with highest Whittle's index $W_k(m_k)$ are activated. Further discussion on the Lagrangian relaxation approach, the indexability property and the definition and derivation of Whittle's index can be found in Chapter 2.

It has been shown that the Whittle index policy performs strikingly well, see Niño-Mora [72] for a discussion, and is asymptotically optimal under certain conditions, see Ji *et al.* [60], Ouyang *et al.* [74], Verloop [93] and Weber *et al.* [96]. The latter explains the importance given in the literature to the calculation of Whittle's index. In addition to resource allocation problems, Whittle's index has been applied in a wide variety of cases, including opportunistic scheduling in wireless downlink channels Aalto *et al.* [1], Taboada *et al.* [86, 87], website morphing and pharmaceutical trials, Gittins *et al.* [48, Chapter 9]. In Avrachenkov *et al.* [11] Whittle's index policy was applied to schedule a crawler to retrieve ephemeral content and in Niño-Mora *et al.* [73] to obtain hunting policies for hiding targets. Many other problems have motivated this approach, see for instance Ansell *et al.* [5], Liu *et al.* [66, 67], Raghunathan *et al.* [78] and Singh *et al.* [85]. The recent survey paper Glazebrook *et al.* [52] is a good reference on the application of index policies in scheduling.

1.3.2 Fluid approximation

The approach that we are going to introduce in this section consists in approximating the original stochastic process by a deterministic system, by only taking into account the mean drifts. In the context of this thesis we will perform this approximation in RBPs with birth-and-death transitions. Let us denote the birth rates of bandit k by $b_k^a(m) := q_k^a(m, m+1)$ for all $a \in \mathcal{A}_k$ and $m \in E_k$, and the death rates by $d_k^a(m) := q_k^a(m, m-1)$, where $a \in \mathcal{A}_k$ and $m \in \{1, 2, \ldots\}$. Define $m_k(t) \in \mathbb{R}^+ \cup \{0\}$ as the amount of *fluid* in bandit k for all $k \in \{1, \ldots, K\}$. Then, we approximate the stochastic process $(N_1(t), \ldots, N_K(t))$ by $(m_1(t), \ldots, m_K(t))$ where $m_k(t)$ is described by

$$\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = b_k^a(m_k(t)) - d_k^a(m_k(t)),$$

for all $k \in \{1, \ldots, K\}$. We shall assume that $b_k^a(m_k)$ and $d_k^a(m_k)$ are continuous in $m_k \in \mathbb{R}^+ \cup \{0\}$. Also, we consider the cost function $C_k(m_k, a)$ for all $k \in \{1, \ldots, K\}$ to be continuous in $m_k \in \mathbb{R}^+ \cup \{0\}$.

The deterministic model that comes out from this approach falls in the framework of optimal control theory, and it is easier to solve compared to the original stochastic model. Standard techniques to find optimal solutions for deterministic control problems are presented in Section 1.3.4. The deterministic control problem may lose some of the properties of the original optimization problem but may provide important insights.

We expect that the deterministic process $(m_1(t), \ldots, m_K(t))$ can be obtained as a result of a fluid scaling. For example speeding up the time, while scaling the original process itself accordingly, one might expect to converge to a deterministic limiting process. In the thesis we present $m_k(t)$ for all $k \in \{1, \ldots, K\}$ as a mere approximation of the original process $N_k(t)$ and we provide no guarantee on whether it is a result of some fluid-scaled process.

The approach of using the fluid model to find an approximation for the stochastic model finds its roots in the pioneering works by Avram *et al.* [14] and Weiss [97]. It is remarkable that in some cases the optimal control for the fluid model coincides with the optimal solution for the stochastic problem. See for example Avram *et al.* [14] where this is shown for the $c\mu$-rule in a multi-class single-server queue and Bäuerle *et al.* [18] where this is shown for Klimov's rule in a multi-class queue with feedback. In some other cases the optimal fluid solution does not coincide with the optimal stochastic control but serves as a heuristic, see for instance Avram *et al.* [12, 13], Brun *et al.* [32]. In that case, asymptotical optimality results may hold, that is, that the fluid-based control is optimal for the stochastic optimization problem after a suitable scaling, see for example Baüerle *et al.* [17], Gajrat *et al.* [44], Maglaras [68], Meyn *et al.* [69], Verloop *et al.* [94]. We conclude by mentioning that the fluid approach owes its popularity to the groundbreaking result stating that if the fluid model drains in finite time, the stochastic process is stable, see Dai [37], Robert [79, Chapter 9].

1.3.3 Optimal stochastic control

To solve the optimization of RBPs we will make use of optimal stochastic control techniques. We will explain them in this section for a general model, *i.e.,* we consider cost per unit of time $C(m, a)$ and the transition rates $q^a(m, \tilde{m})$ for all $m, \tilde{m} \in E$ and $a \in \mathcal{A}$, where E is the state space and $\mathcal{A}$ the action space.

We assume $E \subseteq \mathbb{N}^d$, for some $d \in \mathbb{N}$. The objective is to minimize the long-run average cost:

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T C(N^\phi(t), S^\phi(N^\phi(t)))\mathrm{d}t\right). \tag{1.3.4}$$

Under mild assumptions, one can prove that for this type of models a pure stationary policy is average optimal [54]. Therefore, one can focus on stationary policies. However, finding an optimal decision policy is far from being trivial and the analysis has to be done in a case-by-case basis.

We assume that decisions are made at transition times. Since the time in between two transitions follows an exponential distribution, we can convert the continuous-time MDP to an equivalent discrete-time MDP. This adaptation can be done using the uniformization technique. To do so we assume that the transition time in state m, which is exponentially distributed with parameter $\tau(m,a) := \sum_{\tilde{m}\in E} q^a(m,\tilde{m})$, is uniformly bounded, *i.e.,* $\tau(m,a) \leq b$ for all states $m \in E$ and all actions $a \in \mathcal{A}$. Sufficient conditions for an optimal solution to exist can be obtained as the solution of the optimality equations. There exist $\tilde{g}$ and $V(\cdot)$ such that

$$\tilde{g} + \tau(m,a)V(m) = \min_{a\in\mathcal{A}}\{C(m,a) + \sum_{\tilde{m}\in E} q^a(m,\tilde{m})V(\tilde{m})\}, \tag{1.3.5}$$

where $\tilde{g}$ is the per unit of time average optimal cost and $V(\cdot)$, which is known as the value function, captures the difference in cost between starting in state m and an arbitrary reference state. Multiplying the latter by $1/b$ we obtain

$$\frac{\tilde{g}}{b} + V(m) = \min_{a\in\mathcal{A}}\left\{\frac{C(m,a)}{b} + \sum_{\tilde{m}\in E}\frac{q^a(m,\tilde{m})}{b}(V(\tilde{m}) - V(m)) + V(m)\right\}.$$

Recall that $\sum_{\tilde{m}\in E} q^a(m,\tilde{m}) = \tau(m,a)$, and defining $g := \tilde{g}/b$, $p^a(m,\tilde{m}) = \frac{q^a(m,\tilde{m})}{b}$ the transition probabilities for all m and $\tilde{C}(m,a) = C(m,a)/b$ the cost per transition, we obtain the following optimality equation

$$g + V(m) = \min_{a\in\mathcal{A}}\{\tilde{C}(m,a) + \sum_{\tilde{m}\in E} p^a(m,\tilde{m})V(\tilde{m})\}. \tag{1.3.6}$$

The value g can now be interpreted as the optimal average cost per transition. Equation (1.3.6) is also known as Bellman equation and is the discrete-time analogous of the original MDP.

When the transition rates $\tau(m,a)$ are not uniformly bounded, the uniformization method can not be performed. In order to overcome this issue, one can truncate the system, that is, we select $L \in E$ such that the transitions to states above L are not allowed. If in the truncated state space we do have bounded transition rates, then the uniformization can be performed. This is the method used in Figure 1.3. In Figure 1.3 (left) we have used a small truncating parameter L, whereas in Figure 1.3 (right) we have considered a large L. We observe that using a small truncation parameter makes the so-called boundary effect to appear, which moves up as the parameter L increases. Analytically, one might be interested in proving structural properties of the value function, such as, convexity/concavity, monotonicity or supermodularity/subadditivity, which might be needed in order to characterize optimal policies. In this context the truncation technique explained above is not a good approach, since structural properties might

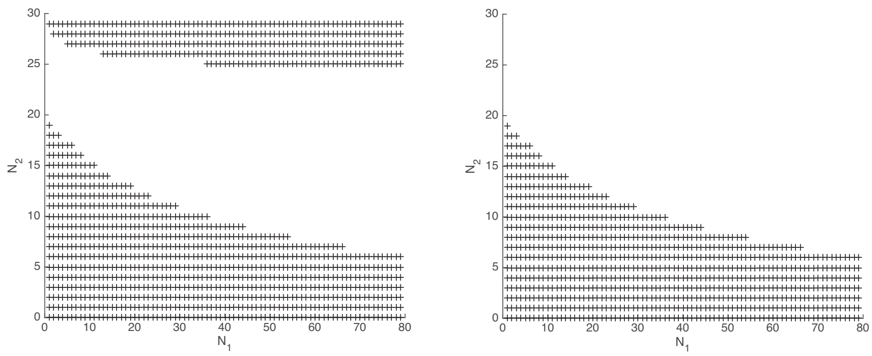


Figure 1.3: Value iteration solution for a two-class single-server abandonment queue. In the area with "+" class-1 customers are prioritized, and in the blanc area class-2 customers. The figure in the left shows a boundary effect for truncation parameters $L_1 = 80, L_2 = 30$. The figure in the right shows a solution by value iteration for large truncation parameters ($L_1 = L_2 = 100$), for which the boundary effect moved up.

be lost due to boundary effects. The authors in Bhulai *et al.* [25] propose the Smoothed Rate Truncation (SRT) method, which consists in truncating the infinite state space with a parameter L and linearly decreasing the relevant transitions such that $q^a(m, \ell) = 0$ for all $\ell \in \{m \in E : \exists i \in \{1, \dots, d\}, m_i > L_i\}$ and $m \in E$. Under certain conditions it is proven that the value function of the smoothed system V^L, satisfies $V^L \to V$ as $L \to \infty$ and has the same structural properties.

Optimality equation (1.3.6) can be solved using a Dynamic Programming (DP) technique, we refer to Bellman [19], Bertsekas [22, Chapter 1] and Ross [80, Chapters II, III and IV] for an understanding on this technique. We are going to present the value iteration algorithm which is one of the two DP algorithms used to solve MDPs and which provides an optimal stationary policy. Let $\epsilon > 0$,

- *Step 0*: select $V_0(m) = 0$ for all $m \in E$.
- *Step 1*: define
$$V_{t+1}(m) = \min_{a \in \mathcal{A}} \{\tilde{C}(m, a) + \sum_{\tilde{m} \in E} p^a(m, \tilde{m}) V_t(\tilde{m})\}.$$
- *Step 2*: if
$$\max_{m \in E}\{V_{t+1}(m) - V_t(m)\} - \min_{m \in E}\{V_{t+1}(m) - V_t(m)\} \le \epsilon,$$
we set $\varphi(m) = \arg\min_{a \in \mathcal{A}}\{\tilde{C}(m, a) + \sum_{\tilde{m} \in E} p^a(m, \tilde{m}) V_t(\tilde{m})\}$. Otherwise augment t by $t + 1$ and go to *Step 1*.

In order for the value iteration algorithm to converge certain extra conditions have to be satisfied Puterman [77, Chapter 8.5]. Nevertheless, if it does converge then one obtains that $V_t(m) \to V(m)$ as $t \to \infty$, $V_{t+1}(m) - V_t(m) \to g$, and $\varphi(m)$ converges to an optimal pure stationary policy.

The value iteration algorithm will be exploited in this thesis for two main purposes. Firstly, to numerically obtain an approximation to the average optimal policy. This will allow us to numerically compare the heuristics derived throughout the thesis with respect to an optimal solution that comes out as a result of value iteration. Secondly, we will use the value iteration algorithm to prove optimality of threshold policies, where threshold policies are such that above a threshold you do action a and below the threshold action a' with $a \neq a' \in \mathcal{A}$. The latter will be done in Section 3.2.1 and Section 6.3. Under certain properties of the value function one can prove for instance that threshold type of policies are an optimal solution of the optimization problem.

1.3.4 Optimal deterministic control

In this section we consider the problem of finding an optimal dynamic control for a deterministic system whose dynamics are described by differential equations. This problem falls in the framework of optimal control theory. Optimal control theory has been for many years used to solve problems in mathematics and engineering.

Below we introduce the mathematical formulation of an optimal control problem and we set the necessary and sufficient conditions for optimality, see also Appendix A. More detailed statements and proofs can be found in Bertsekas [22, Vol I, Chapter 3], Chachuat [34, Chapter 3], Hartl *et al.* [56], Treat *et al.* [90, Part II].

Formulation of an optimal control problem

In an optimal control problem we encounter two classes of variables, the *state variables* and the *control variables*, the latter modifies the former. The objective is to find a control trajectory $s(t)$ for all $t \in [0, T]$ together with the state trajectory $m(t)$ such that the functional

$$\int_0^T C(m(t), s(t))\,\mathrm{d}t, \tag{1.3.7}$$

is minimized, where the cost function $C(\cdot,\cdot)$ is continuously differentiable with respect to $m(\cdot)$, continuous with respect to $s(\cdot)$ and T is the terminal time which can either be fixed or free (subject to optimization). We further assume that $m(t) \in \mathbb{R}^d$, for some $d \in \mathbb{N}$, and $s(t) \in \mathcal{S}$, with $\mathcal{S}$ the set of all admissible controls.

The dynamics are given by a set of differential equations, that is,

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} = f(m(t), s(t)), \text{ for all } t \in [0, T], \tag{1.3.8}$$

and the initial state is $m(0) = m_0$.

In optimal control problems all sorts of state and control constraints can be encountered, nevertheless, in the framework of RBP we will only consider the following type of constraints

$$h_1(s(t)) \leq 0, \text{ and } h_2(m(t)) \leq 0, \text{ for all } t \in [0, T], \tag{1.3.9}$$

where $h_1(\cdot)$ is a pure control constraint, *i.e.*, it only depends on the control $s(t)$, and $h_2(\cdot)$ is a pure state constraint, *i.e.*, it only depends on the state $m(t)$. We further assume that $h_2(\cdot)$ is a first order constraint, that is,

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}h_2(m(t))}{\mathrm{d}t}\right) = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}h_2(m(t))}{\mathrm{d}m(t)} f(m(t), s(t))\right) \neq 0. \tag{1.3.10}$$

We call $(m(t), s(t))$ a feasible pair if constraints (A.2.2) are satisfied and $s(t) \in \mathcal{S}$. Also $(m(t), s(t))$ is called an optimal pair if Equation (1.3.7) is globally minimized.

We will discuss the necessary and sufficient conditions for an optimal control problem. We will address Pontryagin's Minimum Principle (PMP), which provides necessary conditions for optimality, and the Hamilton-Jacobi-Bellman (HJB) equation, which is a sufficiency result.

Necessary conditions for optimality

PMP is due to L.S. Pontryagin, a Russian mathematician who did groundbreaking research on control and established the basis for modern optimization theory, see Pontryagin *et al.* [76]. PMP provides necessary conditions for the pair $(m(t), s(t))$ to be optimal. In the case in which the control and the state of the system appear linearly, and the state and control constraints are rather *simple*, PMP can be used as a constructive method to build *extremal solutions*. By extremal solutions we denote all admissible pairs $(m(t), s(t))$ that satisfy the necessary conditions for optimality. Note that, necessary conditions permit only to restrict the set of $(m(t), s(t))$ pairs, but do not provide an optimal solution.

Before introducing the PMP we define the Hamiltonian and the Lagrangian of an optimal control problem. The Hamiltonian is defined as:

$$\mathcal{H}(m(t), s(t), \gamma(t)) := C(m(t), s(t)) + \gamma^{\mathsf{T}}(t) f(m(t), s(t)), \tag{1.3.11}$$

where $\gamma(\cdot)$ is the adjoint vector. In the presence of additional constraints, like those in Equation (A.2.2), we can define the Lagrangian as:

$$\mathcal{L}(m(t), s(t), \gamma(t), \nu(t), \omega(t)) := \mathcal{H}(m(t), s(t), \gamma(t)) + \nu^{\mathsf{T}}(t) h_1(s(t)) + \omega^{\mathsf{T}}(t) h_2(m(t)),$$

where $\nu(\cdot)$ and $\omega(\cdot)$ are Lagrange multipliers.

Having defined the Hamiltonian and the Lagrangian of the optimal control problem, denoted from now on by $\mathcal{H}(t)$ and $\mathcal{L}(t)$, we can now state the necessary conditions for the optimality of the pair $(m(t), s(t))$:

$$\begin{aligned}
&\dot{\gamma}(t) = -\frac{\partial \mathcal{L}(t)}{\partial m}, \quad \frac{\partial \mathcal{L}(t)}{\partial s} = 0, \quad s(t) = \arg \min_{s(t) \in \mathcal{S}} \mathcal{H}(t), \\
&\nu(t) h_1(s(t)) = 0, \nu(t) \geq 0 \text{ and } \omega(t) h_2(m(t)) = 0, \omega(t) \geq 0, \text{ for all } t \in [0, T].
\end{aligned}$$

A formal description can be found in Appendix A.2, at the end of the thesis, and we refer to Hartl *et al.* [56] for a survey on necessary conditions in the presence of state constraints.

The necessary conditions for optimality will be used in Chapter 5 to obtain a solution of the optimal control problem that corresponds to a two-class single-server abandonment queue under the total cost criteria.

Sufficient conditions for optimality

The HJB equation is one of the main results of the optimal control theory. Suppose there exists a function $V(t, m)$ continuously differentiable in t and m such that

$$0 = \min_{s \in \mathcal{S}} [C(m, s) + \nabla_t V(t, m) + \nabla_m V(t, m)^{\mathsf{T}} f(m, s)], \text{ for all } t, m, \tag{1.3.12}$$

where $\nabla_t V(t,m)$ and $\nabla_m V(t,m)$ stand for the partial derivative with respect to time and state, respectively. Then, value function $V(t,m)$ is the optimal cost of the optimal control problem (1.3.7)-(A.2.2), and the control trajectory $s^*(m)$ that minimizes the right-hand-side of (1.3.12) is optimal. A more formal statement can be found in Appendix A.3. Equation (1.3.12) is known as the HJB equation. The statement of the sufficiency theorem and a complete proof can be found in Bertsekas [22, Prop. 3.2.1].

In most cases the sufficient conditions for optimality do not help in finding a solution, since for high dimensional spaces the HJB equation is typically intractable to solve. However, the necessary conditions for optimality can reduce the set of admissible controls considerably and the HJB can then be used to verify the optimality of the so-obtained candidates for optimality.

1.4 Overview of the thesis

In this section we present the main contributions of the thesis. This first introductory chapter has been devoted to motivate the modeling approach as well as to give an overview of the methodology that has been considered.

The remainder of the thesis is structured in three parts. In Part I we present the general framework of RBPs in the case in which the state of the bandits evolves in a birth-and-death fashion. In Chapter 2 we introduce the notions of indexability and Lagrangian relaxation, and we derive Whittle's index in closed-form expression. The expression depends on the steady-state probabilities which are well known for birth-and-death processes. This representation of Whittle's index allows us to define the indexability as a property on the steady-state probabilities. In Chapter 3 we apply the results in Chapter 2 to the particular example of an abandonment queue. We prove that threshold policies are optimal for the unconstrained problem and that bandits are indexable. The latter provides an expression for Whittle's index, which for particular cases, can be computed explicitly. We notice that abandonments can help to recover indices for non-indexable problems letting the abandonment rate tend to 0, this is for instance the case for the M/M/1 queue with convex holding cost. We prove that the Whittle index is asymptotically optimal in the heavy-traffic regime and light-traffic regime, and we numerically evaluate Whittle's index policy's performance across several workloads. We conclude that Whittle's index policy is nearly optimal in the abandonment queue. The results presented in Chapter 2 and Chapter 3 can be found in [SR2] and [SR5].

In Part II we approximate the multi-class stochastic decision making problem by simpler deterministic ones. Two different approaches will be considered. In Chapter 4 we start from the relaxation approach considered in Part I, and scale the relaxed version of the multi-class multi-server problem under the assumption that each queue evolves in a birth-and-death fashion. This gives rise to a deterministic unconstrained problem, where now the Lagrange multiplier W can be tuned to obtain an index policy for the fluid problem, which we will refer to as *fluid index*. Once the fluid index has been derived one can develop the analogous theory of Whittle's index and we then propose the fluid index policy as a heuristic. This new heuristic has several advantages over Whittle's index, namely, it can be explicitely derived, and indexability and monotonicity of the optimal (fluid) control can be easily proven. We show the applicability of this heuristic by considering several case studies that belong to the class of examples introduced in Section 1.2.2. We observe that this new heuristic performs well across several workloads and coincides with Whittle's index in several instances. The results presented in Chapter 4 are based on the papers [SR3] and [SR6]. In Chapter 5 we study the multi-class abandonments model considered in Chapter 3, however, in this case we approximate the original stochastic problem by a fluid model. We then split the analysis into the underload case and the overload case. In the overload case we are able to completely characterize the optimal policy, which turns out to be a strict priority rule. In the underload case we solve the two-class case using PMP and based on this solution we build a heuristic for a system with an arbitrary number of classes. The optimal solution in the two-class system appears to be a threshold type of policy. We observe in extensive numerical experiments that the heuristic proposed in this chapter captures well the structure of the optimal policy for the original stochastic model. The results in Chapter 5 are based on the paper [SR4].

In Part III we further study the phenomena of abandonments in queueing systems. Contrary to the study carried out in the previous parts of the thesis, we now consider the dynamics in the queue to be different from the simple birth-and-death structure, namely, we allow batch departures to occur. This is motivated by the content delivery problem, where requests for a content might be grouped to be transmitted in multi-cast transmission mode. We analyze the multi-cast transmission mode for a single-class infinite server capacity system. In this chapter we prove that threshold type of policies are average optimal and we derive the queue-length distribution in two different settings: (1) when the service rate is infinite, and (2) when the service rate is finite. This allows us to characterize the optimal transmission policy (which is of threshold type). We observe numerically the importance of computing the optimal threshold due to the poor performance of non-optimal threshold policies. Part III is based on the paper [SR1].

Finally in Appendix A, we present a compilation of relevant optimality results for stochastic and deterministic control problems.

Part I

Dynamic control of stochastic resource-sharing systems

Chapter 2

Index policy for birth-and-death restless bandits

Contents

2.1 Introduction **19**
2.2 Model description **20**
2.3 Relaxation and indexability **22**
2.4 Whittle's index policy **27**
2.5 Performance of Whittle's index policy **27**
2.6 Appendix **28**

In this chapter, we consider the framework of RBP introduced in Section 1.2.1 in the particular case in which bandits evolve in a birth-and-death fashion. We derive Whittle's index policy and we give sufficient conditions for the indexability property (required to derive Whittle's index) to hold. This heuristic allows us to define a decision control for complex resource allocation problems. At the end of the chapter we propose some steps to prove asymptotic optimality of Whittle's index policy for limiting regimes. The framework developed in this chapter will be later exploited in Chapter 3 and Chapter 4 to derive heuristics to efficiently share resources in several systems, *e.g.*, multi-class abandonment queues, wireless downlink channels, power-aware server farms and inventory management with perishable items.

2.1 Introduction

An optimal policy to a RBP will in general be a complex function of all the input parameters and the number of competing bandits. In practice optimal policies can be found only for very specific instances. A particular case in which an optimal solution can be found, as discussed in Section 1.2.1, is the MABP context for which the so-called *index policy* is optimal. Optimality of index policies has enjoyed a great popularity. The solution to a complex control problem that, a priori, might depend on the entire state space, turns out to have a strikingly simple structure. A classical example is a multi-class single-server queue with linear holding costs where it is known that the celebrated $c\mu$ index rule is optimal, that is, it is optimal to serve the classes in decreasing order of priority according to the product $c_k\mu_k$, where c_k is the holding cost per class-k customer, and μ_k^{-1} is the mean service requirement of class-k customers, Buyukkoc *et al.* [33], Gelenbe *et al.* [46]. The simple structure of the optimal policy vanishes however in the presence of, *e.g.*, convex costs, servers with state-dependent capacities and/or impatient customers Ansell *et al.* [5], Bispo [26] and George *et al.* [47]. Another classical result that can be seen as an index policy is the optimality of Shortest-Remaining-Processing-Time (SRPT), where the index of each customer is given by its remaining service time Schrage *et al.* [83].

In the more general RBP case, obtaining optimal solution is out of reach. We will therefore look for well-performing heuristics, such as Whittle's index policy. In general Whittle's index has to be derived in a case-by-case basis. In this chapter we will show that in the particular case in which bandits evolve in a birth-and-death fashion, provided that some extra properties such as the monotonicity of optimal policies as well as indexability hold, Whittle's index can be obtained in closed-form expression. Birth-and-death processes have many applications in demography, queueing theory, performance engineering, epidemiology and biology and will therefore be the focus of this chapter.

The remainder of the chapter is organized as follows. In Section 2.2, we introduce the model under study, in Section 2.3, we further discuss the Lagrangian relaxation method for birth-and-death bandit evolution, we present threshold type of policies and the notion of indexability. The latter two properties allow us to obtain a closed-form expression of Whittle's index. In Section 2.4, we define Whittle's index policy and finally, in Section 2.5, we discuss a framework in order to prove asymptotic optimality of Whittle's index in limiting regimes. All the proofs can be found in Section 2.6.

2.2 Model description

We consider a stochastic resource allocation problem with K bandits. Let $N_k(t) \in \{0, 1, \dots\}$ denote the state of bandit k at time t, $k = 1, \dots, K$. Decision epochs are defined as the moments when a bandit changes state. At each decision epoch, the scheduler can choose for each bandit between two actions: action $a = 0$, that is, making the bandit passive, or action $a = 1$, that is, making the bandit active, with the restriction that at any moment in time at most M bandits can be made active. Throughout this chapter we consider bandits that are modeled as a continuous time birth-and-death process, that is, when bandit k is in state m_k, it changes the state after an exponentially distributed amount of time, and can go either to state $(m_k - 1)^+$ or state $m_k + 1$. The transition rates for bandit k depend only on m_k (and not on the state of the other bandits). More precisely, when N_k denotes the state of bandit $k = 1, \dots, K$, the transition rates of the vector $\vec{m} = (m_1, \dots, m_K)$ are

$$\begin{cases} \vec{m} \to \vec{m} + \vec{e}_k & \text{with transition rate} \quad b_k^a(m_k), \\ \vec{m} \to \vec{m} - \vec{e}_k & \text{with transition rate} \quad d_k^a(m_k), \end{cases} \tag{2.2.1}$$

where $\vec{e}_k$ is a K-dimensional vector with all zeros except for the k-th component which is equal to 1, and $d_k^a(0) = 0$.

We note that the transitions of a bandit depend on the action chosen. In particular, the state of bandits can evolve both when being active and passive.

A policy ϕ decides which bandit is made active. Because of the Markov property, we can focus on policies that base their decisions only on the current state of the bandits. For a given policy ϕ, $N_k^\phi(t)$

denotes the state of bandit k at time t and $\vec{N}^\phi(t) = (N_1^\phi(t), \ldots, N_K^\phi(t))$. Let $S_k^\phi(\vec{N}^\phi(t)) \in \{0, 1\}$ represent whether or not bandit k is made active at time t under policy ϕ. At most M out of the K bandits can be made active, or equivalently, at least $K - M$ bandits have to be passive. Hence, we have the constraint

$$\sum_{k=1}^{K} (1 - S_k^\phi(\vec{N})) \geq K - M, \tag{2.2.2}$$

which is equivalent to (1.2.2). For bandit k, let $C_k(m, a)$ denote the cost per unit of time when in state m and it is either passive (action $a = 0$) or active (action $a = 1$) and assume $C_k(\cdot, \cdot)$ to be convex non-decreasing and bounded by a polynomial of finite degree. We denote by $\mathcal{U}$ the set of stationary ergodic control policies that satisfy constraint (2.2.2).

The objective is to find a scheduling policy, denoted by OPT, that minimizes the long-run expected average-cost criteria

$$\mathcal{C}^\phi := \limsup_{T\to\infty} \sum_{k=1}^{K} \frac{1}{T} \mathbb{E}\left(\int_0^T C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t)))\ \mathrm{d}t \right). \tag{2.2.3}$$

We denote by $\mathcal{C}^{OPT} := \min_{\phi\in\mathcal{U}} \mathcal{C}^\phi$ the average cost under the optimal policy OPT. For problem (2.2.3), it is known that if there exist g and $V(\cdot)$ that satisfy the Bellman equation

$$\begin{aligned} g = \min_{\vec{s}, s.t. \sum_k s_k \leq M} \Bigg(\sum_{k=1}^{K} \Big[& C_k(m_k, s_k) + b_k^{s_k}(m_k) V(\vec{m} + e_k) + d_k^{s_k}(m_k) V(\vec{m} - e_k) \\ & - (d_k^{s_k}(m_k) + b_k^{s_k}(m_k)) V(\vec{m}) \Big] \Bigg), \end{aligned} \tag{2.2.4}$$

a stationary policy that realizes the minimum in (2.2.4) is optimal, see Section 1.3.3. Equation (2.2.4) is the analogous to Equation (1.3.5) for the particular case in which K bandits evolve in a birth-and-death fashion. Here, $g = \min_\phi \mathcal{C}^\phi$ and $V(\vec{m})$ is the value function.

In the next section we will motivate the approach we have used in this chapter by introducing two classes of examples that fall in the RBP birth-and-death context.

2.2.1 Examples

Our main motivation to study this problem, as briefly presented in Section 1.2.2, comes from resource allocation problems arising in multi-class multi-server environments. Assuming there are K classes of customers, each class is represented by a bandit, and the state N_k of bandit k represents the number of class-k customers in the system. Furthermore, $b_k^a(m_k)$ and $d_k^a(m_k)$ denote the arrival and departure rate, respectively. Having a state-dependent departure rate allows us to model important phenomena such as power-aware server farms and user impatience in which customers may leave the system before finishing service. In the former the departure rate will be proportional to the speed-scaling term $(m_k)^\alpha$, see Wierman *et al.* [100], and in the latter the departure rates will include a term $\theta_k m_k$, where θ_k is the abandonment rate of class-k customers, see Atar *et al.* [8] and Glazebrook *et al.* [49]. To illustrate the applicability of our framework, we now present two broad classes of problems that fall inside the RBP framework. Both classes of examples are further developed in Chapter 3 and Chapter 4 (Sec. 4.3).

The first class of problems concerns the multi-class setting of Figure 2.1 (left). The objective is to determine which M classes to be made simultaneously active. Hence, the transition rates are as follows: $b_k^a(m_k) = \lambda_k(m_k)$ and $d_k^a(m_k) = \mu_k(m_k, a)$, where $a = 1$ in case class k is served. We enable the arrival rate of each class to depend on its queue length and the departure rate to depend on its queue length and the action $a \in \{0, 1\}$. In Chapter 3 and Section 4.3.1 we use this model to study optimal scheduling in a multi-class single-server queue with customers abandonment and in Section 4.3.2 to study optimal scheduling in a wireless downlink problem where, as a consequence of opportunistic scheduling, the capacity increases with the number of customers, see Borst [28]. We further note that when $M = 1$ and $\mu_k(m_k, a) = \mu_k$, this model captures the classical single-server multi-class queue.

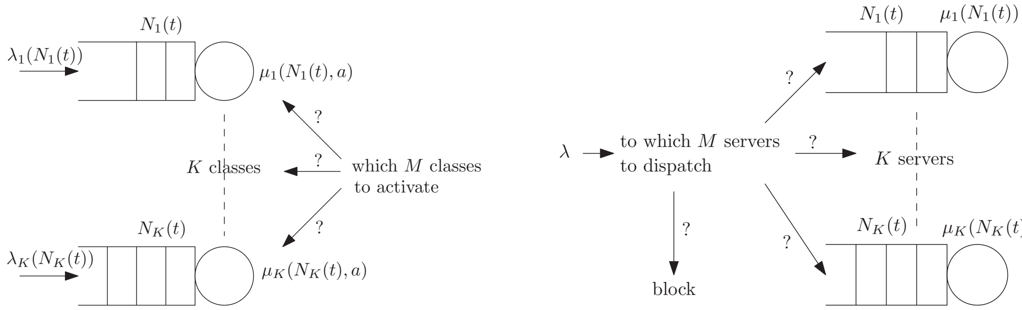


Figure 2.1: Left: A multi-class system where M classes can be simultaneously served. Right: Load balancing in a multi-server system

The second class of problems is the load balancing problem, see Figure 2.1 (right), where new arrivals must be dispatched to K heterogeneous servers, or must be blocked. We allow an arrival to be dispatched to at most M servers (simultaneously), where $M = 1$ is the typical value for load-balancing problems. Hence, the transition rates are as follows: $b_k^a(m_k) = \lambda a$ and $d_k^a(m_k) = \mu_k(m_k)$, where $a = 1$ in case an arrival is routed to server k. In Sections 4.3.3 and 4.3.4 we investigate how to optimally dispatch customers in (i) a power-aware server farm, where the capacity of servers follows a speed-scaling rule and in (ii) a make-to-stock production system, where stocked items are perishable.

2.3 Relaxation and indexability

In this section we will describe the Lagrangian relaxation method, which was proposed by Whittle [99] and which we briefly discussed in Section 1.3.1, to derive approximate solutions to (2.2.3) under constraint (2.2.2). A fruitful approach has been to study the relaxed problem in which the constraint on the number of active bandits must be satisfied on average, and not in every decision epoch, that is,

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \sum_{k=1}^K (1 - S_k^\phi(\vec{N}^\phi(t)))\mathrm{d}t\right) \geq K - M. \tag{2.3.1}$$

The objective of the relaxed problem is hence to determine the policy that solves (2.2.3) under constraint (2.3.1). An optimal policy for the relaxed problem, which turns out to be of index type, then serves as heuristic for the original optimization problem. We denote by $\mathcal{U}^{REL}$ the set of policies that satisfy (2.3.1) are stationary and for all $\phi \in \mathcal{U}^{REL}$ the Markov chain is ergodic. We note that $\mathcal{U} \subset \mathcal{U}^{REL}$.

The relaxed problem can be solved by considering the following unconstrained problem: find a policy ϕ that minimizes

$$\mathcal{C}^{\phi}(W)$$
$$:= \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\bigg(\int_0^T \bigg(\sum_{k=1}^{K} C_k(N_k^{\phi}(t), S_k^{\phi}(\vec{N}^{\phi}(t))) + W(K - M - \sum_{k=1}^{K}(1 - S_k^{\phi}(\vec{N}^{\phi}(t)))\bigg)\mathrm{d}t\bigg), \qquad (2.3.2)$$

where W is the Lagrange multiplier. For a given W, let $REL(W)$ denote a policy that minimizes (2.3.2), and let $\mathcal{C}^{REL(W)}(W) := \min_{\phi\in\mathcal{U}^{REL}} \mathcal{C}^{\phi}(W)$ denote the optimal performance of the relaxed problem. For any value of the multiplier $W \geq 0$, it holds that $\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^{OPT}$. To see this note that for a given $W \geq 0$ and $\phi \in \mathcal{U}$ it holds that

$$\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^{\phi}(W) \leq \mathcal{C}^{\phi}.$$

The first inequality follows by definition of $REL(W)$, and the second inequality follows from the fact that $M - K + \sum_{k=1}^{K}(1 - S_k^{\phi}(\vec{N}^{\phi}(t))) \geq 0$ for any policy $\phi \in \mathcal{U}$.

Problem (2.3.2) can be decomposed into K subproblems, one for each different bandit k, which under the stationarity and ergodicity assumption simplifies to the following: find a policy ϕ that minimizes

$$\mathcal{C}_k^{\phi}(W) := \mathbb{E}(C_k(N_k^{\phi}, S_k^{\phi}(N_k^{\phi}))) - W\mathbb{E}(\mathbf{1}_{S_k^{\phi}(N_k^{\phi})=0}), \qquad (2.3.3)$$

where N_k^{ϕ} is distributed as the stationary distribution of the state of bandit k under policy ϕ. The solution to (2.3.2) is obtained by combining the solution to the K separate optimization problems (2.3.3).

Problem (2.3.3) is a MDP as well and the optimal policy is the solution of the Bellman equation

$$\begin{aligned} g_k(W) = \min \Big(& C_k(m,1) + b_k^1(m)\Delta V(m) - d_k^1(m)\Delta V(m-1), \\ & C_k(m,0) - W + b_k^0(m)\Delta V(m) - d_k^0(m)\Delta V(m-1)\Big), \end{aligned} \qquad (2.3.4)$$

with $g_k(W) = \min_{\phi} \mathcal{C}_k^{\phi}(W)$ the minimum cost under an optimal policy, and $\Delta V(m) = V(m+1) - V(m)$.

In some particular instances the solution to the problem (2.3.3) can be proven to be of threshold type. In the next section we introduce the definition of threshold and some of the scenarios in which it can be established.

2.3.1 Threshold policies

For certain problems, it can be established that the structure of an optimal solution of problem (2.3.3) is of threshold type. That is, it can be shown that: there is a threshold $n_k(W)$ such that when bandit k is in a state $m_k \leq n_k(W)$, then action a is optimal, and otherwise action a' is optimal, $a, a' \in \{0, 1\}$ and $a \neq a'$. We let policy $\phi = n$ denote a threshold policy with threshold n, and we refer to it as 0-1 type if $a = 0$ and $a' = 1$, and 1-0 type if $a = 1$ and $a' = 0$ (see Figure 2.2). Optimality of a threshold policy for an unconstrained optimization problem in the RBP context, has been proved for example in Ansell *et al.* [5], Glazebrook *et al.* [49] and Veatch *et al.* [91]. Further examples can be found in Gittins *et al.* [48, Section 6.5].

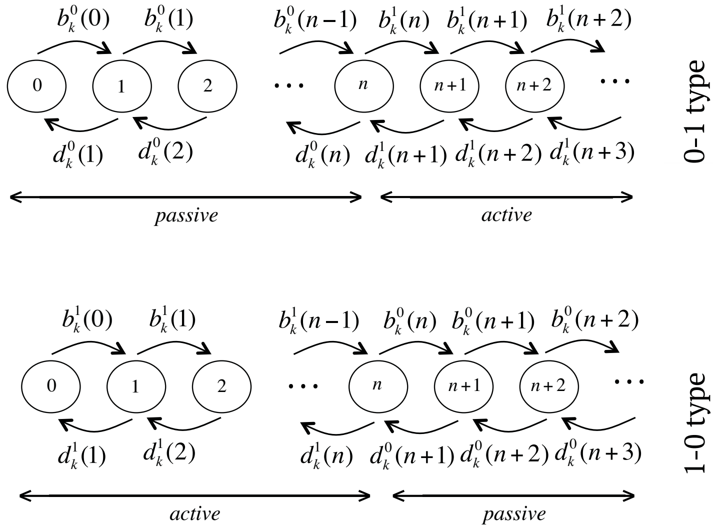


Figure 2.2: Transitions under threshold policies; above the 0-1 type, below the 1-0 type.

In the next proposition we give sufficient conditions for an optimal policy that solves problem (2.3.3) to be of threshold type. The proof can be found in Appendix 2.6.1.

Proposition 2.1. *Assume $b_k^a(m_k) = \lambda_k(m_k)$ and $d_k^a(m_k) = \mu_k(m_k)a$. Assume there exists $\phi \in \mathcal{U}^{REL}$. Then there exists an $n_k = -1, 0, 1, \ldots$ such that a 0-1 type of threshold policy, with threshold n_k, optimally solves problem* (2.3.3)*. If instead, $b_k^a(m_k) = \lambda_k(m_k)a$ and $d_k^a(m_k) = \mu_k(m_k)$. Then there exists an $n_k = -1, 0, 1, \ldots$ such that a 1-0 type of threshold policy, with threshold n_k, optimally solves problem* (2.3.3)*.*

Outside the framework assumed in Proposition 2.1, no sufficient conditions are known that imply monotonicity for a general bandit with birth-and-death state evolution. We refer to Chapter 3 where threshold type of policies have been proved to be optimal for a multi-class abandonment queue.

In the next section we present the definitions of indexability and Whittle's index. Moreover, in the case in which threshold type of policies can be proven to be optimal the Whittle index will be derived.

2.3.2 Indexability and Whittle's index

Indexability is the property that allows us to develop a heuristic for the original problem. This property requires to establish that as the Lagrange multiplier, or equivalently the subsidy for passivity, W, increases, the collection of states in which the optimal action is *passive* increases. It was first introduced by Whittle [99] and we formalize it in the following definition.

Definition 2.1. *A bandit is indexable if the set of states in which passive is an optimal action in* (2.3.3) *(denoted by $D_k(W)$) increases in W, that is, $W' < W \Rightarrow D_k(W') \subseteq D_k(W)$.*

If indexability is satisfied, Whittle's index in state N_k is defined as follows:

Definition 2.2. *When a bandit is indexable, Whittle's index in state m_k is defined as the smallest value for the subsidy such that an optimal policy for* (2.3.3) *is indifferent of the action in state m_k. The Whittle's index is denoted by $W_k(m_k)$.*

The solution to the relaxed control problem (2.3.2) will be to activate all bandits that are in a state n_k such that their Whittle's index exceeds the subsidy for passivity, *i.e.*, $W_k(n_k) > W$. In particular, a standard Lagrangian argument shows that there exists a value $W = W^*$, for which the constraint (2.3.1) is binding, *i.e.*, the optimal policy ϕ that solves problem (2.3.2) for $W = W^*$ will on average activate (at most) M bandits.

We are now in position to derive Whittle's index when an optimal policy for problem (2.3.3) is fully characterized by a threshold n. The average cost can be expressed as a function of the steady-state probabilities, which in the case of birth-and-death processes has a well-known solution. We let $\pi_k^n(m)$ denote the steady-state probability of bandit k of being in state m under threshold policy n. Let us define

$$g_k^{(n)}(W) := \sum_{m=0}^{\infty} C_k(m,a)\pi_k^n(m) - \begin{cases} W\sum_{m=0}^{n}\pi_k^n(m) & \text{if threshold policy is of 0-1 type,} \\ W\sum_{m=n+1}^{\infty}\pi_k^n(m) & \text{if threshold policy is of 1-0 type.} \end{cases}$$

Then the optimal average cost for problem (2.3.3) is given by

$$g_k(W) = \min_n \{g_k^{(n)}(W)\}. \tag{2.3.5}$$

We now state the steps to obtain Whittle's index. The proof of Proposition 2.2 can be found in Appendix 2.6.2.

Proposition 2.2. *Assume an optimal solution of* (2.3.3) *is of threshold type, and* $\sum_{i=0}^{n}\pi_k^n(i)$ *is strictly increasing in* n*, with* $\pi_k^n(m)$ *the steady-state probability of being in state* m *under threshold policy* n*. Then problem* (2.3.3) *is indexable and Whittle's index values for bandit* k *are computed by the following steps:*

- **Step 0** *Compute*

$$W_0 = \inf_{n\in\mathbb{N}\cup\{0\}} \frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{-1}, S_k^{-1}(N_k^{-1})))}{\sum_{m=0}^{n}\pi_k^n(m)},$$

  and name by n_0 *the largest minimizer if the threshold policies are of 0-1 type. Compute*

$$W_0 = \sup_{n\in\mathbb{N}\cup\{0\}} -\frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{-1}, S_k^{-1}(N_k^{-1})))}{\sum_{m=0}^{n}\pi_k^n(m)},$$

  and name by n_0 *the largest maximizer if the threshold policies are of 1-0 type. Then, define* $W_k(n) := W_0$ *for all* $n \le n_0$*. If* $n_0 = \infty$ *define* $W_k(n) := W_0$ *for all* $n > n_0$*, otherwise go to Step 1.*

- **Step j** *Compute*

$$W_j = \inf_{n\in\mathbb{N}\setminus\{0,\ldots,n_{j-1}\}} \frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{n_{j-1}}, S_k^{n_{j-1}}(N_k^{n_{j-1}})))}{\sum_{m=0}^{n}\pi_k^n(m) - \sum_{m=0}^{n_{j-1}}\pi_k^{n_{j-1}}(m)}, j \ge 1,$$

  and name by n_j *the largest minimizer if the threshold policies are of 0-1 type. And compute*

$$W_j = \sup_{n\in\mathbb{N}\setminus\{0,\ldots,n_{j-1}\}} -\frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{n_{j-1}}, S_k^{n_{j-1}}(N_k^{n_{j-1}})))}{\sum_{m=0}^{n}\pi_k^n(m) - \sum_{m=0}^{n_{j-1}}\pi_k^{n_{j-1}}(m)}, j \ge 1,$$

and name by n_j the largest maximizer if the threshold policies are of 1-0 type. Then, define $W_k(n) := W_j$ for all $n_{j-1} < n \leq n_j$. If $n_j = \infty$ then $W_k(n) = W_j$ for all $n > n_j$, otherwise go to step $j+1$.

Proposition 2.2 gives a general recipe to develop Whittle's indices for bandits whose evolution can be described by general birth-and-death processes:

(i) Establish optimality of threshold policies.
(ii) Establish indexability, that is, prove that $\sum_{m=0}^{n} \pi_k^n(m)$ is strictly increasing.
(iii) If (i) and (ii) can be established, then Whittle's index is given by Proposition 2.2.

Steps (i) and (ii) are model dependent. Step (iii) is immediate and the index will always be given by Proposition 2.2. The latter procedure will be followed in Chapter 3 for a multi-class abandonment queue, and in Chapter 4 in several load balancing and optimal class selection problems.

In the next corollary we characterize the Whittle index in the particular case in which $n_i = i$ for all $i \in \mathbb{N} \cup \{0\}$, with n_i as defined in Proposition 2.2. The proof can be found in Appendix 2.6.3

Corollary 2.1. *Assume an optimal solution of* (2.3.3) *is of threshold type, and $\sum_{m=0}^{n} \pi_k^n(m)$ is strictly increasing in n, with $\pi_k^n(m)$ the steady-state probability for bandit k of being in state m under threshold policy n. Then, bandit k is indexable.*

If the structure of an optimal solution of problem (2.3.3) is of 0-1 type, then, in case

$$\frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{n-1}, S_k^{n-1}(N_k^{n-1})))}{\sum_{m=0}^{n} \pi_k^n(m) - \sum_{m=0}^{n-1} \pi_k^{n-1}(m)}, \tag{2.3.6}$$

is non-decreasing in n, Whittle's index $W_k(n_k)$ is given by (2.3.6) *and is hence non-decreasing. Similarly, if the structure of an optimal solution of problem (2.3.3) is of 1-0 type, then, in case* (2.3.6) *is non-decreasing in n, $-W_k(n_k)$ is given by* (2.3.6) *and hence Whittle's index is non-increasing.*

We shall now make a comment which concerns the form of (2.3.6). The numerator in (2.3.6) can be interpreted as the increase in cost by deciding to become passive in state n and keeping all other actions unchanged, and similarly, the denominator can be understood as the corresponding increase of passivity rate for the process, measured by the additional probability in which a subsidy is received. Thus, $W(n)$ can be interpreted as a measure of increased cost per unit of increased passivity, a term coined as Marginal Productivity Index by Niño-Mora [72].

To the best of our knowledge, it has not been reported previously that for bandits whose evolution can be described by a birth-and-death process, one can get an explicit closed-form expression for Whittle's index. Perhaps a reason for this lies in the difficulty to solve the optimality equation (2.3.4), which has two unknowns g_k and $V_k(m)$. This has led researchers to circumvent this difficulty by considering the discounted cost first, whose Bellman's equation has only one unknown, equating the total discounted costs as done in Proposition 2.2 for average cost and then taking the limit in order to retrieve an index for the average cost case. This is for instance the approach taken in Ansell *et al.* [5] or in Gittins *et al.* [48, Section 6.5] for bi-directional bandits in which the active and passive actions push the process in opposite directions. In Glazebrook *et al.* [49] the authors develop an algorithm to calculate an index in a multi-class queue with admission control. All these models have in common that after the relaxation, the bandits are birth-and-death, and the obtained Whittle's index is thus equal to (2.3.6). Regarding the bi-directional bandit it can be directly checked that index (2.3.6) is equivalent to the index of Gittins *et al.* [48, Theorem 6.4].

2.4 Whittle's index policy

In this section we describe how the optimal solution to the relaxed optimization problem is used to obtain a heuristic for the original model. The optimal solution to the relaxed problem, that is, activate all bandits that are in a state n_k such that $W_k(n_k) > W$, might be unfeasible for the original model where at most M bandits can be served at a time. Hence, Whittle [99] proposed the following heuristic, which is referred to as Whittle's index policy.

Definition 2.3 (Whittle's index policy)**.** *Assume at time t we are in state $\vec{N}(t) = \vec{n}$. The Whittle index policy activates the M bandits having currently the highest non-negative Whittle's index value $W_k(n_k)$.*

Note that in case all bandits are in a state such that their Whittle's index is negative, all bandits are kept passive. The latter is a direct consequence of the relaxed optimization problem: when the Whittle index is negative for a bandit in state $\tilde{n}$, this means that it is made active only if $W < W_k(\tilde{n}) < 0$, that is, when a *cost* is payed for being passive.

2.5 Performance of Whittle's index policy

Having defined the Whittle index policy in the previous section, we can now discuss the possibility of proving near optimality of this heuristic. This discussion will be exploited in Section 3.5.2 in order to prove optimality of Whittle's index policy in limiting regimes, such as light-traffic and heavy-traffic regimes.

We recall that $\mathcal{U}$ and $\mathcal{U}^{REL}$ refer to the set of admissible policies in the original and relaxed problem, respectively, and that $\mathcal{U} \subseteq \mathcal{U}^{REL}$. As we argued in Section 2.2, for any value of the multiplier $W \geq 0$, $\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^{OPT}$, where $\mathcal{C}^{REL(W)}(W)$ and $\mathcal{C}^{OPT}$ are the minimum cost in the relaxed and original problems, respectively. We also recall that $\mathcal{C}^{REL(W)}(W)$ is achieved by a policy that serves all the classes with current Whittle's index larger than W. We denote by $\mathcal{C}^{WI}$ the performance in the original problem under *the admissible* Whittle's index policy and we set $\mathcal{C}^* = \sup_W \mathcal{C}^{REL(W)}(W)$. It then trivially holds that

$$\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^* \leq \mathcal{C}^{OPT} \leq \mathcal{C}^{WI}. \tag{2.5.1}$$

We now argue that if either

(i) $REL(0) \in \mathcal{U}$, or,

(ii) $REL(W) \in \mathcal{U}$ and the constraint (2.3.1) is satisfied with equality,

then it holds that, for that choice of W, $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^* = \mathcal{C}^{OPT} = \mathcal{C}^{WI}$, and hence in those cases Whittle's index policy is optimal for the original policy. This can be seen as follows. First we observe that if $REL(W) \in \mathcal{U}$, then $REL(W)$ coincides with Whittle's index policy. Hence, for $W = 0$ we have $\mathcal{C}^{REL(0)}(0) = \mathcal{C}^{REL(0)} = \mathcal{C}^{WI}$, where the first equality holds by definition since $W = 0$. Now assume $W > 0$, then since (2.3.1) holds with equality, we have again $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^{REL(W)} = \mathcal{C}^{WI}$. In both cases, we use (2.5.1) to conclude that $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^* = \mathcal{C}^{OPT} = \mathcal{C}^{WI}$. We note that the same approach is described in Gittins *et al.* [48, Chapter 6] and Glazebrook *et al.* [49, Section 5].

When either (i) or (ii) can be proved to hold for limiting regimes, asymptotic optimality of Whittle's index policy follows directly from it.

2.6 Appendix

2.6.1 Proof of Proposition 2.1

Let us drop the dependency on k throughout the proof.

Let us consider the case $b^a(m) = \lambda(m)$ and $d^a(m) = \mu(m)a$, the case $b^a(m) = \lambda(m)a$ and $d^a(m) = \mu(m)$ can be done similarly. Since there exist $\phi \in \mathcal{U}^{REL}$, then there exists an stationary policy ϕ^* that optimally solves problem (2.3.3). Define $n^* = \max\{m \in \{0, 1, \ldots\} : S^{\phi^*}(m) = 0\}$, then note from the definition of the transition rates that all states $m < n^*$ are transient. Hence $\pi^{\phi^*}(m) = 0$ for all $m < n^*$. Moreover, by definition of n^* we have $S^{\phi^*}(n^*) = 0$ and $S^{\phi^*}(m) = 1$ for all $m > n^*$. The average cost as given by (2.3.3) under the optimal policy ϕ^* then reduces to

$$\mathbb{E}(C(N^{\phi^*}, S^{\phi^*}(N^{\phi^*}))) - W\mathbb{E}(\mathbf{1}_{S^{\phi^*}(N^{\phi^*})=0}) = C(n^*, 0)\pi^{\phi^*}(n^*) + \sum_{m=n^*+1}^{\infty} C(m, 1)\pi^{\phi^*}(m) - W\pi^{\phi^*}(n^*)$$
$$= \mathbb{E}(C(N^{n^*}, S^{n^*}(N^{n^*}))) - W\pi^{n^*}(n^*),$$

that is, a 0-1 type of threshold policy with threshold n^* gives optimal performance.

2.6.2 Proof of Proposition 2.2

We will focus on 0-1 type of threshold policies throughout the proof, the case in which threshold n is of 1-0 type can be proven similarly.

Let us assume that the steps stop at iteration $J \in \mathbb{N} \cup \{\infty\}$, and hence $n_J = \infty$. We further set $W_i := W_J$ and $n_i = \infty$ for all $i \in \{J+1, \ldots\} \cup \{\infty\}$. We will prove that $W_0 < W_1 < W_2 < \ldots < W_J$, and note that by construction n_i for $i \in \mathbb{N} \cup \{0, \infty\}$ is an increasing sequence. Let us prove $W_i < W_{i+1}$ for all $i \in \{0, 1, 2, \ldots, J\}$. We have from the characterization of W_i that

$$\frac{\mathbb{E}(C(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)}$$
$$> \frac{\mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)}$$
$$\Longrightarrow (\mathbb{E}(C(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))) \left(\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)\right)$$
$$> (\mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))) \left(\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)\right),$$

where the latter inequality follows from $\sum_{m=0}^{n} \pi^n(m)$ being strictly increasing. Adding the term

$$\mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i}))(\sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m) - \sum_{m=0}^{n_i} \pi^{n_i}(m)),$$

on both sides of the inequality, after some algebra we obtain

$$\begin{aligned} W_{i+1} &= \frac{\mathbb{E}(C(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i})))}{\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_i} \pi^{n_i}(m)} \\ &> \frac{\mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)} = W_i. \end{aligned}$$

Observe that $W_{i+1} > W_i$ together with $n_{i+1} > n_i$ imply indexability.

To prove that the steps given in Proposition 2.2 indeed define the Whittle index we have to show that,

1. the threshold policy -1 is optimal for problem (2.3.3) for all W such that $W < W_0$.
2. The threshold policy $n_i < \infty$ is optimal for problem (2.3.3) for all W such that $W_i < W < W_{i+1}$.
3. And finally that the policy ∞, is optimal for problem (2.3.3) for all W such that $\infty > W > W_J$ and $J < \infty$.

To show 1., note that for all $W < W_0$

$$\begin{aligned} & W \sum_{m=0}^{n} \pi^n(m) < \mathbb{E}(C(N^n, S^n(N^n))) - \mathbb{E}(C(N^{-1}, S^{-1}(N^{-1}))), \\ & \Longrightarrow \mathbb{E}(C(N^{-1}, S^{-1}(N^{-1}))) < \mathbb{E}(C(N^n, S^n(N^n))) - W \sum_{m=0}^{n} \pi^n(m), \text{ for all } n \geq 0, \end{aligned}$$

that is, $g^{(-1)}(W) < g^{(n)}(W)$ for all $n \in \mathbb{N} \cup \{0\}$, and hence $g(W) = g^{(-1)}(W)$. Policy -1 is therefore optimal for problem (2.3.3) for $W < W_0$.

We will prove 2. by induction. Observe from the definition of n_0 that for all $n \geq 0$

$$\mathbb{E}(C(N^{n_0}, S^{n_0}(N^{n_0}))) - W_0 \sum_{m=0}^{n_0} \pi^{n_0}(m) \leq \mathbb{E}(C(N^n, S^n(N^n))) - W_0 \sum_{m=0}^{n} \pi^n(m), \tag{2.6.1}$$

that is, $g^{(n_0)}(W_0) \leq g^{(n)}(W_0)$, for all $n \geq 0$. Besides, we trivially have that $g^{(n_0)}(W_0) \leq g^{(-1)}(W_0)$. By the assumption that $\sum_{m=0}^{n} \pi^n(m)$ is strictly increasing in n, and we obtain from (2.6.1) that

$$\begin{aligned} & \mathbb{E}(C(N^{n_0}, S^{n_0}(N^{n_0}))) - W \sum_{m=0}^{n_0} \pi^{n_0}(m) \leq \mathbb{E}(C(N^n, S^n(N^n))) - W \sum_{m=0}^{n} \pi^n(m) \\ & \Longrightarrow g^{(n_0)}(W) \leq g^{(n)}(W), \end{aligned}$$

for all $n \leq n_0$ and $W_0 < W$. In particular, $g^{(n_0)}(W) \leq g^{(n)}(W)$ is satisfied for all $W_0 < W < W_1$ and $n \leq n_0$. Similarly, from the definition of W_1 we have that $g^{(n_0)}(W_1) \leq g^{(n)}(W_1)$ for all $n \geq n_0 + 1$, and again using that $\sum_{m=0}^{n} \pi^n(m)$ is strictly increasing we obtain $g^{(n_0)}(W) \leq g^{(n)}(W)$ for all $W_0 < W < W_1$ and $n \geq n_0 + 1$.

We have therefore proven that $g^{(n_0)}(W) \leq g^{(n)}(W)$ for all n and $W_0 < W < W_1$, that is, policy n_0 is optimal for all W such that $W_0 < W < W_1$. This establishes the first step of the induction $i = 0$. Let us now assume that it holds for step $i - 1 \geq 0$, that is, n_i is an optimal policy for problem (2.3.3), given W such that $W_{i-1} < W < W_i$. And let us assume $n_i < \infty$. The definition of W_i together with the fact that

n_{i-1} is optimal for the choice of W such that $W_{i-1} < W < W_i$, imply

$$g^{(n_{i-1})}(W_i) = g^{(n_i)}(W_i) \leq g^{(n)}(W_i), \text{ for all } n \geq 0.$$

Recall that $\sum_{m=0}^{n} \pi^n(m)$ is strictly increasing in n and therefore

$$g^{(n_i)}(W) \leq g^{(n)}(W), \text{ for all } n \leq n_i, \text{ and all } W \text{ such that } W_i < W < W_{i+1}.$$

Besides, from the definition of W_{i+1} we have

$$g^{(n_i)}(W) \leq g^{(n)}(W), \text{ for all } n \geq n_i + 1, \text{ and all } W \text{ such that } W_i < W < W_{i+1}.$$

We therefore have obtained that threshold policy n_i is optimal for problem (2.3.3) given W such that $W_i < W < W_{i+1}$.

Finally, we prove 3. for $J < \infty$, note that from the induction followed in the previous point we have that

$$g^{(n_{J-1})}(W_J) = g^{(n_J)}(W_J) \leq g^{(n)}(W_J), \text{ for all } n \geq 0,$$

and the fact that $\sum_{m=0}^{n} \pi^n(m)$ is increasing in n gives that

$$g^{(n_J)}(W) < g^{(n)}(W), n \leq n_J = \infty, \text{ for all } W \text{ such that } W_J < W.$$

Which concludes the proof of the theorem.

2.6.3 Proof of Corollary 2.1

For ease of notation, we removed the subscript k. Since an optimal policy for (2.3.3) is of threshold type, for a given subsidy W the optimal average cost under threshold n will be $g(W) := \min_n\{g^{(n)}(W)\}$, where for 0-1 type

$$g^{(n)}(W) := \sum_{m=0}^{\infty} C(m, S^n(m))\pi^n(m) - W \sum_{m=0}^{n} \pi^n(m), \tag{2.6.2}$$

and for 1-0 type

$$\begin{aligned} g^{(n)}(W) &:= \sum_{m=0}^{\infty} C(m, S^n(m))\pi^n(m) - W \sum_{m=n+1}^{\infty} \pi^n(m) \\ &= \sum_{m=0}^{\infty} C(m, S^n(m))\pi^n(m) + W \sum_{m=0}^{n} \pi^n(m) - W. \end{aligned} \tag{2.6.3}$$

We denote by $n(W)$ the minimizer of $g(W)$.

The function $g(W)$ is a lower envelope of affine non-increasing functions in W. It thus follows that $g(W)$ is a concave non-increasing function. See Figure 2.3 for an illustration, where we plot $g^{(n)}(W)$ for the case in which n is of 0-1 type on the left and for the case 1-0 type on the right.

It follows directly that the right-derivative of $g(W)$ in W is given by $-\sum_{m=0}^{n(W)} \pi^{n(W)}(m)$ for the 0-1 structure ($\sum_{m=0}^{n(W)} \pi^{n(W)}(m) - 1$ for the 1-0 structure). Since $g(W)$ is concave in W, the right derivative is

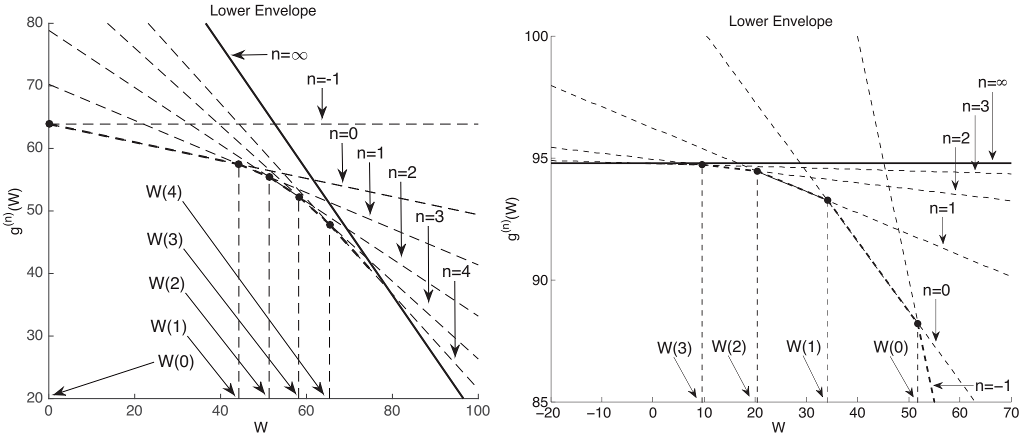


Figure 2.3: We plot the lower envelop *i.e.,* $g = \min_n\{g^{(n)}\}$. Left: A multi-class abandonment model where $b^a(m) = \lambda$, $d^a(m) = \mu a + \theta m$, $C(m,a) = (1+2\theta)m + 3m^2$, for $a = 0,1$, and $\theta = 6, \lambda = 23, \mu = 10$. Right: A make-to-stock problem with perishable items where $b^a(m) = \mu a$, $d^a(m) = \lambda + \theta m$, $C(m,a) = (2+3\theta)m + m^2 + 14\pi^m(0)$, and $\mu = 5, \lambda = 10, \theta = 2.5$.

non-increasing in W. Together with the fact that $\sum_{m=0}^{n} \pi^n(m)$ is strictly increasing in n, it hence follows that $n(W)$ is non-decreasing (non-increasing) in W. Since an optimal policy is of threshold type, the set of states where it is optimal to be passive can be written as $D(W) = \{m : m \le n(W)\}$ for the 0-1 type (or $D(W) = \{m : m \ge n(W)\}$ for the 1-0 type). Since $n(W)$ is non-decreasing (non-increasing), by definition this implies that bandit k is indexable.

Let $\tilde{W}(n)$ be the value for the subsidy such that the average cost under threshold policy n is equal to that under policy $n-1$. Hence, using (2.3.3), we have that for all $n \ge 1$, $\mathbb{E}(C(N^n, S^n(N^n)) - \tilde{W}(n)\mathbb{E}(\mathbb{1}_{S^n(N^n)=0})$ is equal to $\mathbb{E}(C(N^{n-1}, S^{n-1}(N^{n-1})) - \tilde{W}(n)\mathbb{E}(\mathbb{1}_{S^{n-1}(N^{n-1})=0})$. When the threshold policy n is of type 0-1, then $\mathbb{E}(\mathbb{1}_{S^n(N^n)=0}) = \sum_{m=0}^{n} \pi^n(m)$, hence $\tilde{W}(n)$ is given by (2.3.6). When the threshold policy n is of type 1-0, then $\mathbb{E}(\mathbb{1}_{S^n(N^n)=0}) = \sum_{m=n+1}^{\infty} \pi^n(m)$, hence $\tilde{W}(n)$ is given by

$$\begin{aligned}
&\frac{\mathbb{E}(C(N^n, S^n(N^n))) - \mathbb{E}(C(N^{n-1}, S^{n-1}(N^{n-1})))}{\sum_{m=n+1}^{\infty} \pi^n(m) - \sum_{m=n+1}^{\infty} \pi^{n-1}(m)} \\
&= -\frac{\mathbb{E}(C(N^n, S^n(N^n))) - \mathbb{E}(C(N^{n-1}, S^{n-1}(N^{n-1})))}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} = -(2.3.6).
\end{aligned}$$

It can be verified that $\tilde{W}(n)$ being monotone, implies that $g(\tilde{W}(n)) = g^{(n)}(\tilde{W}(n)) = g^{(n-1)}(\tilde{W}(n))$. In addition, for 0-1 type of threshold policies, since $\frac{\mathrm{d}g^{(n)}(W)}{\mathrm{d}W} = -\sum_{m=0}^{n} \pi^n(m)$ is decreasing in n, we have $g(W) = g^{(n-1)}(W)$ for $\tilde{W}(n-1) \le W \le \tilde{W}(n)$. This implies that Whittle's index is given by $W(n) = \tilde{W}(n)$. Similarly, for 1-0 type of policies, $\frac{\mathrm{d}g^{(n)}(W)}{\mathrm{d}W} = -\sum_{m=n+1}^{\infty} \pi^n(m)$ is increasing in n, and then $g(W) = g^{(n)}(W)$ for $\tilde{W}(n) \le W \le \tilde{W}(n-1)$. This implies that Whittle's index is given by $W(n) = \tilde{W}(n)$.

Chapter 3

Index policy for an abandonment queue

Contents

3.1 Model description and preliminaries 34
3.2 Relaxation 36
3.3 Case studies 39
3.4 Multi-class M/M/1 queue 43
3.5 Asymptotic optimality 45
3.6 Numerical results 46
3.7 Appendix 51

In this chapter we investigate a resource allocation problem in a multi-class single-server setting, with convex holding costs and customers abandonment under the average cost criterion. This problem fits the framework introduced in Chapter 2. Following the structure presented there, we will relax the optimization problem, prove threshold policies to be optimal for the unconstrained model, establish indexability and derive Whittle's index. The remainder of the chapter will be devoted to deriving expressions and properties of the Whittle index for a particular choice of the input parameters or in limiting regimes. In the case of linear holding cost, we show that the Whittle's index is a constant that is independent of the number of customers in the system and of the arrival rate. For general convex holding cost we derive properties of the index value in limiting regimes: we consider the behavior of the index (i) as the number of customers in a class grows large, which allows us to derive the asymptotic structure of the index policies, (ii) as the abandonment rate vanishes, which allows us to retrieve an index policy proposed for the multi-class M/M/1 queue with convex holding cost and no abandonments, and (iii) as the arrival rate goes to either 0 or ∞, representing light-traffic and heavy-traffic regimes, respectively. Moreover, we will prove that Whittle index policy is optimal in heavy-traffic and light-traffic regimes and we numerically show that it performs very well across all workloads.

The chapter is organized as follows. In Section 3.1 we describe the multi-class single-server abandonment model. In Section 3.2 we apply the Lagrangian relaxation method and we show that threshold policies are optimal for the resulting K unidimensional unconstraint optimization problems. We establish indexability and calculate Whittle's index under the average cost criterion. In Section 3.3 we calculate Whittle's index for linear holding cost and derive properties for general convex holding costs in several limiting regimes. In Section 3.4 we calculate the index for an $M/M/1$ queue without abandonments. Section 3.5 describes our asymptotic optimality results. Finally, in Section 3.6 we numerically evaluate the performance of Whittle's index policy. Most of the proofs are presented in the Appendix 3.7.1.

3.1 Model description and preliminaries

We consider a multi-class single-server queue with K classes of customers. Class-k customers arrive according to a Poisson process with rate λ_k and have an exponentially distributed service requirement with mean $1/\mu_k$, $k = 1, \dots, K$. We denote by $\rho_k := \lambda_k/\mu_k$ the traffic load of class k, and by $\rho := \sum_{k=1}^{K} \rho_k$ the total load to the system. We model abandonments of customers in the following way:

- Any class-k customer not served abandons after an exponentially distributed amount of time with mean $1/\theta_k$, $k = 1, \dots, K$, with $\theta_k > 0$.
- A class-k customer that is being served abandons after an exponentially distributed amount of time with mean $1/\theta'_k$, $k = 1, \dots, K$, with $\theta'_k \geq 0$.

The server has capacity 1 and can serve at most one customer at a time, where the service can be preemptive. We make the following natural assumption:

$$\mu_k + \theta'_k \geq \theta_k, \text{ for all } k.$$

That is, the departure rate of a class-k customer is higher when being served than when not being served.

At each moment in time, a policy ϕ decides which class is served. Because of the Markov property, we can focus on policies that only base their decisions on the current number of customers present in the various classes. For a given policy ϕ, $N_k^\phi(t)$ denotes the number of class-k customers in the system at time t, (hence, including the one in service), and $\vec{N}^\phi(t) = (N_1^\phi(t), \dots, N_K^\phi(t))$. Let $S_k^\phi(\vec{n}) \in \{0, 1\}$ represent the service capacity devoted to class-k customers at time t under policy ϕ in state $\vec{N}(t) = \vec{n}$. The constraint on the service amount devoted to each class is $S_k^\phi(\vec{n}) = 0$ if $n_k = 0$ and

$$\sum_{k=1}^{K} S_k^\phi(\vec{n}) \leq 1, \tag{3.1.1}$$

and we denote by $\mathcal{U}$ the set of admissible control strategies that satisfy this constraint and we assume that all $\phi \in \mathcal{U}$ are ergodic and the Markov chain under ϕ is ergodic.

The above describes a birth-and-death RBP that makes transitions

$$\vec{n} \to \vec{n} + \vec{e}_k \text{ with rate } \lambda_k, \text{ and,}$$

$$\vec{n} \to \vec{n} - \vec{e}_k \text{ with rate } \mu_k S_k^\phi(\vec{n}) + \theta_k(n_k - S_k^\phi(\vec{n})) + \theta'_k S_k^\phi(\vec{n}),$$

for $n_k > 0$, with $\vec{e}_k$ a K-dimensional vector with all zeros except for the k-th component which is equal to 1.

Let $C_k(n,a)$ denote the cost per unit of time when there are n class-k customers in the system and when either class k is not served (if $a=0$), or when class k is served (if $a=1$). We assume $C_k(\cdot,0)$ and $C_k(\cdot,1)$ are convex, non-decreasing and bounded by polynomials of finite degree and satisfy

$$C_k(n,0)-C_k((n-1)^+,0)\le C_k(n+1,1)-C_k(n,1)\le C_k(n+1,0)-C_k(n,0), \tag{3.1.2}$$

for all $n\ge 0$. Observe that if $C_k(0,0)\ge C_k(0,1)$, then (3.1.2) implies that, for all n, $C_k(n,0)\ge C_k(n,1)$. We also note that (3.1.2) is always satisfied when (i) $C_k(n,a)=C_k(n)$, or when (ii) $C_k(n,a)=C_k((n-a)^+)$. Case (i) represents holding costs for customers in the *system*, while (ii) represents holding costs for customers in the *queue*.

We further introduce a cost δ_k for every class-k customer that abandons the system when not being served and a cost δ'_k for a class-k customer that abandons the system while being served.

The objective of the optimization is to find the optimal scheduling policy, OPT, under the long-run expected average-cost criteria, that is, find the policy ϕ that minimizes

$$\mathcal{C}^\phi:=\limsup_{T\to\infty}\sum_{k=1}^K\frac{1}{T}\mathbb{E}\Big(\int_0^T C_k(N_k^\phi(t),S_k^\phi(\vec{N}^\phi(t)))\ \mathrm{d}t+\delta_k R_k^\phi(T)+\delta'_k R_k'^\phi(T)\Big), \tag{3.1.3}$$

where $R_k^\phi(T)$ and $R_k'^\phi(T)$ denote the number of class-k customers that abandoned the queue while waiting and while being served, respectively, in the interval $[0,T]$ under policy ϕ. We denote by $\mathcal{C}^{OPT}=\inf_{\phi\in\mathcal{U}}\mathcal{C}^\phi$ the average cost under the optimal policy.

We have

$$\mathbb{E}(R_k^\phi(T))=\theta_k\mathbb{E}\left(\int_0^T(N_k^\phi(t)-S_k^\phi(\vec{N}^\phi(t)))\mathrm{d}t\right)$$

and

$$\mathbb{E}(R_k'^\phi(T))=\theta'_k\mathbb{E}\left(\int_0^T S_k^\phi(\vec{N}^\phi(t))\mathrm{d}t\right),$$

by Dynkin's formula Anderson [3, Chapter 6.5]. We introduce the following notation:

$$\tilde{C}_k(n_k,a):=C_k(n_k,a)+\delta_k\theta_k(n_k-a)^++\delta'_k\theta'_k\min(a,n_k), a\in\{0,1\} \tag{3.1.4}$$

so that the objective (3.1.3) can be equivalently written as

$$\limsup_{T\to\infty}\sum_{k=1}^K\frac{1}{T}\mathbb{E}\left(\int_0^T\tilde{C}_k(N_k^\phi(t),S_k^\phi(\vec{N}^\phi(t)))\ \mathrm{d}t\right). \tag{3.1.5}$$

The above described stochastic control problems have proved to be very difficult to solve. Already for the special case of linear holding cost, deriving structural properties of optimal policies is extremely challenging. For example, in Down *et al.* [41] optimal dynamic scheduling is studied for two classes of customers (K = 2), with $\delta_k=\delta'_k$, $\theta_k=\theta'_k$, $\mu_1=\mu_2=1$, and linear holding cost, $C_k(n,a)=c_kn$. Define $\tilde{c}_k:=c_k+\delta_k\mu_k$. For the special case where $\tilde{c}_1\ge\tilde{c}_2$ and $\theta_1\le\theta_2$, the authors show that it is optimal to give strict priority to class 1, see Down *et al.* [41, Theorem 3.5]. It is intuitively clear that giving priority to class 1 is the optimal thing to do, since serving class 1 myopically minimizes the (holding and abandonment) cost and in addition it is advantageous to keep the maximum number of class-2 customers in the system (without idling), since they have the highest abandonment rate. In Bhulai *et al.* [24] optimal dynamic scheduling is studied for $C_k(n,a) = c_k n$, $\delta_k = \delta'_k$, and either $\theta_k = \theta'_k$ or $\theta'_k = 0$. For the special case where the classes can be ordered such that $\tilde{c}_1 \geq \cdots \geq \tilde{c}_K$, $\tilde{c}_1(\mu_1 + \theta'_1 - \theta_1) \geq \cdots \geq \tilde{c}_K(\mu_K + \theta_K - \theta'_K)$, and $\tilde{c}_1(\mu_1 + \theta'_1 - \theta_1)/\theta_1 \geq \ldots \geq \tilde{c}_K(\mu_K + \theta'_K - \theta_K)/\theta_K$, the authors show that it is optimal to give strict priority according to the ordering $1 > 2 > \ldots > K$.

Outside these special parameter settings, or for convex holding cost, an optimal policy is expected to be state dependent, and as far as the authors are aware, no (structural) results exist for this stochastic optimal control problem.

In order to obtain insights into optimal control for convex holding cost, in this chapter we will solve a relaxed version of the optimization problem, as described in Chapter 2. The latter allows us to propose a heuristic for the original model, which we will prove to be optimal in light-traffic and heavy-traffic regimes.

3.2 Relaxation

As presented in Chapter 2 the Lagrangian relaxation technique, under the ergodicity assumption, reduces the optimization problem to the following: find ϕ such that

$$\mathcal{C}_k^{\phi}(W) = \mathbb{E}(\tilde{C}_k(N_k^{\phi}, S_k^{\phi}(N_k^{\phi}))) - W\mathbb{E}(\mathbf{1}_{\{S_k^{\phi}(N_k^{\phi})=0\}}), \tag{3.2.1}$$

is minimized for each bandit k, where W as explained in Section 2.3 can be seen as the subsidy for passivity. We further define $g_k(W) := \min_{\phi} \mathcal{C}_k^{\phi}(W)$.

In Section 3.2.1 we prove that 0-1 type of threshold policies are optimal for (3.2.1), in Section 3.2.2 we prove that all classes are indexable and in Section 3.2.3 we derive Whittle's index in some cases of interest.

3.2.1 Threshold policies

In the following proposition we show that an optimal solution of the unconstrained problem (3.2.1) is of threshold type, *i.e.*, when the number of customers is above a certain threshold n, the class is served, and not served otherwise. We denote by $\phi = n$, $n = -1, 0, 1, \ldots$, the threshold policy with threshold n, that is, $S_k^n(m) = 1$ if $m > n$, and $S_k^n(m) = 0$ otherwise.

Proposition 3.1. *There is an $n = -1, 0, 1, \ldots$, such that the policy $\phi = n$ is an optimal solution of the unconstrained problem (3.2.1).*

Proof. We drop the dependency on k throughout the proof. The value function $V(m)$ satisfies the Bellman optimality equation for average cost models, see Section 1.3.3, that is,

$$\begin{aligned}(\mu + \theta' + m\theta + \lambda)V(m) + g(W) &= \lambda V(m+1) + \theta(m-1)V((m-1)^+) \\ &+ \min\{\tilde{C}(m,0) - W + (\mu+\theta')V(m) + \theta V((m-1)^+), \tilde{C}(m,1) + (\mu+\theta')V((m-1)^+) + \theta V(m)\},\end{aligned} \tag{3.2.2}$$

where $g(W)$ is the average cost incurred under an optimal policy. Proving optimality of a threshold policy is hence equivalent to showing that if it is optimal in (3.2.2) for state $m+1$, $m \geq 0$ to be passive, then it is also optimal in (3.2.2) for state m to be passive, *i.e.*, $\tilde{C}(m+1,0)-W+(\mu+\theta'-\theta)V(m+1) \leq \tilde{C}(m+1,1)+(\mu+\theta'-\theta)V(m)$, implies $\tilde{C}(m,0)-W+(\mu+\theta'-\theta)V(m) \leq \tilde{C}(m,1)+(\mu+\theta'-\theta)V((m-1)^+)$. A sufficient condition for the above to be true is (3.1.2) together with the inequality $V(m+1)+V((m-1)^+) \geq 2V(m)$, for $m \geq 0$. The latter condition, convexity of the value function, will be proved below, which concludes the proof.

In case of bounded transition rates, one can uniformize the system and use value iteration in order to prove convexity. However, our transition rates are unbounded. Following the SRT method, introduced in Section 1.3.3, we therefore consider the truncated space, truncated by $L > 1$, and smooth the arrival transition rates from m to $m+1$ as follows:

$$q^{\phi,L}(m,m+1) := \lambda \left(1 - \frac{m}{L}\right)^+ = \lambda \max\left(0, 1 - \frac{m}{L}\right),$$

for $m = 0, \ldots, L$, which are linearly decreasing until $q^{\phi,L}(L+1, L) = 0$. Denote by $V^L(m)$ the value function of the L-truncated system. After verifying two conditions, (as done in Appendix 3.7.1), we have by Bhulai *et al.* [25, Theorem 3.1] that $V^L(m) \to V(m)$ as $L \to \infty$. Hence, convexity of the function V is implied by convexity of V^L for all L, and we are left with proving the latter. The latter is uniformizable, hence we can use the value iteration technique in order to prove convexity of V^L. This proof is available in Appendix 3.7.1. $\square$

Below we write the steady-state distribution of threshold policy $\phi = n$. We denote, for each class k, the steady-state probability of being in state i under policy $\phi = n$ by $\pi_k^n(i)$. We have

$$\pi_k^n(i) = \prod_{m=1}^{i} \frac{q_k^n(m-1,m)}{q_k^n(m,m-1)} \pi_k^n(0), \quad i = 1, 2, \ldots, \tag{3.2.3}$$

where $\pi_k^n(0) = \left(1 + \sum_{i=1}^{\infty} \prod_{m=1}^{i} \frac{q_k^n(m-1,m)}{q_k^n(m,m-1)}\right)^{-1}$ and

$$q_k^n(m,m-1) := \begin{cases} \theta_k m & \text{for all } m \leq n, \\ \mu_k + \theta_k' + \theta_k(m-1) & \text{for all } m > n, \end{cases}$$
$$q_k^n(m,m+1) := \lambda_k, \qquad \text{for all } m. \tag{3.2.4}$$

Remark 3.1. *In Proposition 3.1 we established optimality of threshold policies for problem (3.2.1) in the case* $\mu_k+\theta_k' \geq \theta_k$ *and when (3.1.2) is satisfied. If instead* $\mu_k+\theta_k' < \theta_k$*, and in addition* $\tilde{C}_k(m,1) > \tilde{C}_k(m,0)$ *for all* m *(but without requiring (3.1.2) to hold), then (for* $W \geq 0$*) the optimal policy is to be passive in all states* m*. This can be easily seen from Equation* (3.2.2)*, after adding subscript* k*, since being always passive is optimal if for all* m

$$\tilde{C}_k(m,0) - W + (\mu_k + \theta_k' - \theta_k)V_k(m) \leq \tilde{C}_k(m,1) + (\mu_k + \theta_k' - \theta_k)V_k((m-1)^+).$$

The latter follows from the above assumptions and the fact that the value function $V_k(\cdot)$ *is non-decreasing. The proof of* $V_k(\cdot)$ *being a non-decreasing function follows as in Appendix 3.7.1.*

In other cases, we have numerically observed that threshold policies are optimal, but we have not established this formally.

3.2.2 Indexability

In this section we show that for the model under consideration, the classes are indexable.

Proposition 3.2. *All classes are indexable.*

Proof. We drop the dependency on k throughout the proof.

From the proof of Proposition 2.2 we have that indexability is implied by $\sum_{i=0}^{n} \pi^n(i)$ being strictly increasing in n. We therefore prove that $\sum_{i=0}^{n} \pi^n(i)$ is strictly increasing in n, or equivalently, that $1 - \sum_{i=n+1}^{\infty} \pi^n(i)$ is strictly decreasing in n. Using (3.2.3), the latter is equivalent to verifying that

$$\frac{\sum_{m=n+1}^{\infty} \prod_{i=1}^{m} \frac{q^n(i-1,i)}{q^n(i,i-1)}}{\sum_{m=n}^{\infty} \prod_{i=1}^{m} \frac{q^{n-1}(i-1,i)}{q^{n-1}(i,i-1)}} < \frac{1 + \sum_{m=1}^{\infty} \prod_{i=1}^{m} \frac{q^n(i-1,i)}{q^n(i,i-1)}}{1 + \sum_{m=1}^{\infty} \prod_{i=1}^{m} \frac{q^{n-1}(i-1,i)}{q^{n-1}(i,i-1)}}, \tag{3.2.5}$$

holds for all n, where $q^n(\cdot,\cdot)$ are defined in (3.2.4). Note that $q^n(m-1,m) = q^{n-1}(m-1,m)$ for all m and $q^n(m,m-1) = q^{n-1}(m,m-1)$ for all $m \neq n$. From the assumption $\mu + \theta' \geq \theta$ we have $q^n(n,n-1) \leq q^{n-1}(n,n-1)$. Hence, the left-hand-side of (3.2.5) is strictly less than 1, while the right-hand-side is larger than or equal to 1. This proves (3.2.5). □

3.2.3 Whittle's index

Having proven that 0-1 type of threshold policies are optimal and that the model under consideration satisfies the indexability property, we have that the Whittle index can be derived as in Proposition 2.2. In the case in which Equation (2.3.6) can be proven to be non-decreasing in n, then Whittle index is given by (2.3.6).

We could not prove that Whittle's index $W_k(n)$ as given in (2.3.6) is non-decreasing in n in general. However, in many particular cases this property can be established. For instance,

- in the case $\mu_k + \theta'_k = \theta_k$, we have $\pi_k^n(m) = \pi_k^{n-1}(m)$ for all m. Hence (2.3.6) can be written as

$$\frac{\tilde{C}_k(n,0)\pi_k^n(n) - \tilde{C}_k(n,1)\pi_k^{n-1}(n)}{\pi^n(n)} = \tilde{C}_k(n,0) - \tilde{C}_k(n,1). \tag{3.2.6}$$

  By condition (3.1.2) we have that $\tilde{C}_k(n,0) - \tilde{C}_k(n,1) \leq \tilde{C}_k(n+1,0) - \tilde{C}_k(n+1,1)$, hence (2.3.6) is non-decreasing in n. This implies that Whittle's index is given by (3.2.6).
- In Proposition 3.3 it will be proved that when $C_k(n,a)$ is linear in n, (2.3.6) is a constant and therefore non-decreasing in n. Hence, Whittle's index is given by (2.3.6).
- Letting $\theta_k \to 0$, we obtain in Proposition 3.7 that

$$\lim_{\theta_k \to 0} \theta_k \frac{\mathbb{E}(\tilde{C}_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(\tilde{C}_k(N_k^{n-1}, S_k^{n-1}(N_k^{n-1})))}{\sum_{m=0}^{n} \pi_k^n(m) - \sum_{m=0}^{n-1} \pi_k^{n-1}(m)},$$

  is non-decreasing in n.

We now prove that when $\tilde{C}_k(m_k, 0) \geq \tilde{C}_k(m_k, 1)$ for all m_k, the Whittle index $W_k(n_k)$ will always be positive. This can be seen as follows. Recall that $W_k(n_k)$ refers to the value of W such that a threshold policy n_k is an optimal solution of the unconstrained problem. Hence, for all $m_k \leq n_k$, it is optimal to keep the class passive, that is, $\tilde{C}_k(m_k, 0) - W_k(n_k) + (\mu_k + \theta'_k - \theta_k)V(m_k) \leq \tilde{C}_k(m_k, 1) + (\mu_k + \theta'_k - \theta_k)V(m_k - 1)$, as we saw in the proof of Proposition 3.1. Since $\tilde{C}_k(m_k, 0) \geq \tilde{C}_k(m_k, 1)$, $\mu_k + \theta'_k \geq \theta_k$, and $V(\cdot)$ is non-decreasing (see proof of Proposition 3.1), it follows that $W_k(n_k) \geq 0$.

Instead, when $\tilde{C}(m_k, 0) < \tilde{C}(m_k, 1)$ for an m_k, $W_k(n_k)$ can be negative for certain states n_k. For example, when $\theta'_k = \theta_k$ and $\delta'_k \gg \delta_k$. Then, even though the total departure rate of class-k customers is highest when serving class k ($\mu_k + \theta'_k \geq \theta_k$), for certain states n_k it might be better not to serve class k. The latter follows since having a class-k customer abandon while being served, will incur a much higher cost than when it abandons while waiting. Hence, a negative subsidy, that is, a cost, is needed in order for it to be optimal to serve class k.

From the practical point of view, the interest of Whittle's index $W_k(n_k)$ as defined in Proposition 2.2 (after adding subscript k) lies in the fact that the index of class k does not depend on the number of customers present in the other classes j, $j \neq k$. Hence, it provides a systematic way to derive implementable policies which we will show perform very well, see Section 3.6, and are asymptotically optimal in certain settings, see Section 3.5.

3.3 Case studies

In this section we further investigate properties of the Whittle index obtained in Proposition 2.2. In Section 3.3.1 we obtain that the index is state-independent for linear holding cost. In Section 3.3.2 we derive asymptotic properties of the index for general convex holding cost functions.

3.3.1 Linear holding cost

We consider here linear holding cost, that is, $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k \min(n_k, a)$. Hence, under this function, any class-k customer in the queue contributes with c_k to the cost, and a class-k customer in service contributes with c'_k to the cost. In particular, if $c'_k = c_k$, then C_k represents the linear holding cost of customers in the *system* and if $c'_k = 0$ then C_k represents the linear holding cost of customers in the *queue*. These two holding cost functions have been considered in the literature in the context of abandonments, for example Ayesta *et al.* [15] considers the former, while Atar *et al.* [8] takes the latter. From our formula (2.3.6) we will be able to obtain a full characterization of Whittle's index.

It will be convenient to define $\tilde{c}_k := c_k + \delta_k\theta_k$, $k = 1, \dots, K$, which can be interpreted as the total cost per unit of time incurred by a class-k customer in the queue. Similarly, $\tilde{c}'_k := \tilde{c}'_k + \delta'_k\theta'_k$ denotes the total cost per unit of time incurred by a class-k customer in service.

We now present the Whittle index for linear holding cost. The proof can be found in Appendix 3.7.2.

Proposition 3.3. *Assume linear holding cost $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k \min(n_k, a)$. Then, the Whittle index for class k is*

$$W_k(n_k) = \frac{\tilde{c}_k(\mu_k + \theta'_k)}{\theta_k} - \tilde{c}'_k, \text{ for all } n_k. \tag{3.3.1}$$

$W_k(n_k)$	$\theta'_k = \theta_k, \delta'_k = \delta_k$	$\theta'_k = 0$
$c'_k = c_k$	$\frac{\tilde{c}_k\mu_k}{\theta_k}$	$\frac{\tilde{c}_k\mu_k}{\theta_k} - c_k$
$c'_k = 0$	$\frac{\tilde{c}_k\mu_k}{\theta_k} + c_k$	$\frac{\tilde{c}_k\mu_k}{\theta_k}$

Table 3.1: $W_k(n_k)$ for linear holding cost as in Proposition 3.3

An interesting feature of (3.3.1) is that it is independent of the arrival rate λ_k and of the number of class-k customers present, n_k. In Section 3.3.2 we will show that this observation only holds for linear holding costs.

The index (3.3.1) allows for the following interpretation. Consider there is only one class-k customer in the system and no future arrivals, we then have $\tilde{C}_k(1,1) = \tilde{c}'_k$, $\tilde{C}(1,0) = \tilde{c}_k$, $q^1_k(1,0) = \theta_k$, $q^0_k(1,0) = \mu_k+\theta'_k$. Index (3.3.1) can equivalently be written as $(\mu_k + \theta'_k)\left(\frac{\tilde{c}_k}{\theta_k} - \frac{\tilde{c}'_k}{\mu_k+\theta'_k}\right)$, which is equal to

$$q^0_k(1,0)\left(\frac{\tilde{C}(1,0)}{q^1_k(1,0)} - \frac{\tilde{C}_k(1,1)}{q^0_k(1,0)}\right). \tag{3.3.2}$$

Hence, the index can be interpreted as the reduction in cost when making a class-k bandit active instead of keeping him passive (the term within the brackets) during a time lag equal to the departure time in the active phase.

We now consider some particular cases that have been studied in the literature, see also Table 3.1. For example, let us consider first the case in which all customers can abandon the system, *i.e.*, $\theta'_k = \theta_k$, for $k = 1, \ldots, K$, and that the cost for abandonment is the same for both active and passive, so $\delta_k = \delta'_k$. We first assume that all customers in the system incur a holding cost. This implies that $c_k = c'_k$, and thus $\tilde{c}_k = \tilde{c}'_k$. Substituting into (3.3.1) we get $W_k(n_k) = \frac{\tilde{c}_k\mu_k}{\theta_k}$. In the case where only customers in the queue incur a holding cost, *i.e.*, $c'_k = 0$, we have $\tilde{c}_k - \tilde{c}'_k = c_k$, and upon substitution in (3.3.1) we get the index $W_k(n_k) = \frac{\tilde{c}_k\mu_k}{\theta_k} + c_k$.

We now assume that only customers in the queue can abandon, that is, the customer in service will not abandon, hence $\theta'_k = 0$, for $k = 1, \ldots, K$. This is the model assumption of Atar *et al.* [8] and [15]. We first assume that all customers in the system incur a holding cost, that is, $c_k = c'_k$, and we thus get $\tilde{c}'_k = c_k$. From (3.3.1) we get $W_k(n_k) = \frac{\tilde{c}_k\mu_k}{\theta_k} - c_k$. We can similarly calculate the index in the case in which only customers in the queue incur a holding cost, *i.e.*, $c'_k = 0$, to obtain the index $W_k(n_k) = \frac{\tilde{c}_k\mu_k}{\theta_k}$. These two last indices have been derived in Atar *et al.* [8] and Ayesta *et al.* [15], respectively. More specifically, Ayesta *et al.* [15] derives the index $\frac{\tilde{c}_k\mu_k}{\theta_k} - c_k$ when studying one customer and no future arrivals. Interestingly, we observe that the index remains the same in the presence of random arrivals as considered in the model under study in this chapter. When the customer in service does not contribute to the holding cost, our model coincides with that analyzed in Atar *et al.* [8], where it is shown that the index rule $\frac{\tilde{c}_k\mu_k}{\theta_k}$ is asymptotically fluid optimal in a multi-server queue in overload ($\rho > 1$). We therefore conclude that the Whittle's index, that we have derived, retrieves index policies that have been proposed in the literature when studying the system in special parameter regimes.

To finish this subsection we now provide an intuition to understand the result of Proposition 3.3 in the case $\theta'_k = \theta_k$ and $c_k = c'_k$. In this setting, at any moment in time, all customers in the system incur a holding cost c_k and can abandon at rate θ_k. Substituting $\mathbb{E}(\tilde{C}_k(N^{n_k}_k, S^{n_k}_k(N^{n_k}_k))) = \tilde{c}_k\mathbb{E}(N^{n_k}_k)$ and

$W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k}$ in (2.3.6), we get the relation

$$\theta_k(\mathbb{E}(N_k^{n_k-1}) - \mathbb{E}(N_k^{n_k})) = \mu_k \left(\sum_{m=n_k}^{\infty} \pi_k^{n_k-1}(m) - \sum_{m=n_k+1}^{\infty} \pi_k^{n_k}(m) \right),$$

which can be seen as a rate conservation. Indeed, the term on the left-hand-side represents the difference in the average number of customers that abandon the system per time unit when comparing both policies n_k and $n_k - 1$. The right-hand-side represents the difference in the average number of customers that is served per time unit when comparing both policies n_k and $n_k - 1$. The left-hand-side being equal to the right-hand-side is exactly the rate conservation.

3.3.2 Convex holding cost

In this section we characterize Whittle's index, assuming that $W_k(n_k)$ is given by Equation (2.3.6), for general convex non-decreasing holding cost functions. We note that the cost associated to abandonments of customers are linear functions. We can thus use the result of Proposition 3.3 to rewrite Whittle's index as

$$W_k(n_k) = \delta_k(\mu_k + \theta_k') - \delta_k'\theta_k' + W_k^c(n_k), \tag{3.3.3}$$

where

$$W_k^c(n_k) := \frac{\mathbb{E}(C_k(N_k^{n_k}, S^{n_k}(N_k^{n_k}))) - \mathbb{E}(C_k(N_k^{n_k-1}, S^{n_k-1}(N_k^{n_k-1})))}{\sum_{m=0}^{n_k} \pi_k^{n_k}(m) - \sum_{m=0}^{n_k-1} \pi_k^{n_k-1}(m)}$$

is the term corresponding to the holding cost. In the remainder of this section, we will focus on $W_k^c(n_k)$.

First we characterize Whittle's index for large state values. Then we obtain Whittle's index as $\lambda_k \downarrow 0$ and $\lambda_k \uparrow \infty$, representing a light-traffic and heavy-traffic regime, respectively. For all cases, we will observe that for non-linear holding cost Whittle's index is dependent on n_k, that is, is state-dependent.

Whittle's index for large states

Recall that the holding costs $C_k(n_k, 1)$ and $C_k(n_k, 0)$ are upper bounded by polynomials of finite degree. Hence, we can write $C_k(n_k, a) = E_k(n_k, a) + o(1)$, for large values of n_k, where $E_k(n_k, 1) = \sum_{i=0}^{P_k} C_k^{(P_k,i)} n_k^i$, with

$$C_k^{(P_k,i)} := \lim_{n_k \to \infty} \frac{C_k(n_k, 1) - \sum_{j=i+1}^{P_k} C_k^{(P_k,j)} n_k^j}{n_k^i},$$

and $E_k(n_k, 0) = \sum_{i=0}^{Q_k} E_k^{(Q_k,i)} n_k^i$, with

$$E_k^{(Q_k,i)} := \lim_{n_k \to \infty} \frac{C_k(n_k, 0) - \sum_{j=i+1}^{Q_k} E_k^{(Q_k,j)} n_k^j}{n_k^i},$$

with $P_k < \infty$ and $Q_k < \infty$. Due to the cost functions being convex and bounded by polynomials of degree P_k and Q_k we have that $C_k^{(P_k,P_k)} > 0$ and $E_k^{(Q_k,Q_k)} > 0$.

In the following proposition we give the expression for Whittle's index for large states. The proof can be found in Appendix 3.7.3.

Proposition 3.4. *Assume Whittle's index is given as in* (2.3.6). *Let* $C_k(n_k, 1)$ *and* $C_k(n_k, 0)$ *be upper bounded by a polynomial of degree* P_k *and* Q_k *respectively. Then, we have* $W_k(n_k) = W_k^\infty(n_k) + o(1)$*, as* $n_k \to \infty$*, where* $W_k^\infty(n_k) := \delta_k(\mu_k + \theta_k') - \delta_k'\theta_k' + \tilde{W}_k^c(n_k)$ *and*

$$\tilde{W}_k^c(n_k) := (E_k(n_k, 0) - E_k(n_k, 1)) + (\mu_k + \theta_k' - \theta_k)/\theta_k \cdot \left(\sum_{i=1}^{Q_k} E_k^{(Q_k,i)} n_k^{i-1} + \sum_{i=2}^{P_k} C_k^{(P_k,i)} \sum_{j=0}^{i-2} n_k^{i-2-j} \left(\frac{\lambda_k}{\theta_k} \right)^{j+1} \right). \quad (3.3.4)$$

The index $W_k^\infty(n_k)$ *is a non-decreasing function.*

Assume $C_k(n_k, a) = C_k(n_k)$ or $C_k(n_k, a) = C_k((n_k - a)^+)$ with $P_k \geq 2$. In that case, $P_k = Q_k$ and $C_k^{(P_k,P_k)} = E_k^{(Q_k,Q_k)}$. For states that are large enough, the value of $W_k^\infty(n_k)$ is determined by the highest polynomial, which is given by

$$\left(E_k^{(P_k,P_k-1)} - C_k^{(P_k,P_k-1)} + \frac{\mu_k + \theta_k' - \theta_k}{\theta_k} E_k^{(P_k,P_k)} \right) n_k^{P_k-1}. \quad (3.3.5)$$

The latter is independent of the arrival rate λ_k, and hence, so is W_k^∞ for large enough states. This robust index (3.3.5) can serve as an approximation for Whittle's index policy when there are a large number of customers in the system. In Section 3.6 we numerically assess the performance under this index policy $W^\infty(\cdot)$.

Light-traffic indices

We present in the following proposition the expression for Whittle's index as $\lambda_k \downarrow 0$, also referred to as the light-traffic regime. The proof can be found in Appendix 3.7.4. Under the light-traffic assumption, the index can be given in closed form. In Section 3.5 we will use this expression to show that Whittle's index is asymptotically optimal in light traffic.

Proposition 3.5. *Assume Whittle's index* $W_k(n_k)$ *is as given in* (2.3.6). *Then,* $W_k(n_k) = \delta_k(\mu_k + \theta_k') - \delta_k'\theta_k' + W_k^c(n_k)$*, where*

$$\lim_{\lambda_k \downarrow 0} W_k^c(n_k) = C_k(n_k, 0) - C_k(n_k, 1) + (C_k(n_k, 0) - C_k(0, 0)) \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k n_k}.$$

Assuming $C_k(0, 0) = 0$, the above index can be rewritten as follows:

$$\lim_{\lambda_k \downarrow 0} W_k^c(n_k) = (\mu_k + \theta_k' + \theta_k(n_k - 1)) \left(\frac{C_k(n_k, 0)}{\theta_k n_k} - \frac{C_k(n_k, 1)}{\mu_k + \theta_k' + \theta_k(n_k - 1)} \right).$$

This allows us for the following interpretation in light traffic. Given that there are n_k class-k customers, and there are no future arrivals, the index measures the reduction in cost when making a class-k bandit active instead of keeping him passive (the term within the brackets) during a time lag equal to the departure time in the active phase.

Heavy-traffic indices

We present in the following proposition the expression for Whittle's index as $\lambda_k \uparrow \infty$, also referred to as the heavy-traffic regime. The proof can be found in Appendix 3.7.5. Under the heavy-traffic assumption, the index can be given in closed form.

Proposition 3.6. *Assume Whittle's index $W_k(n_k)$ is as given in* (2.3.6)*. Define*

$$W_k^{HT}(n_k) := C_k(n_k, 0) - C_k(n_k, 1) + \frac{\mu_k + \theta_k' - \theta_k}{\theta_k} \frac{\mathbb{E}(C_k(N_k^{n_k-1}, 1))}{\lambda_k/\theta_k},$$

where $N_k^{n_k-1}$ denotes the steady-state number of class-k customers under threshold policy $n_k - 1$, and is defined by the transition rates given in (3.2.4)*. If there exists $z \geq 1$ such that $\frac{\mathbb{E}(C_k(N_k^{n_k-1},1))}{\lambda_k^z} \to 0$, as $\lambda_k \to \infty$, then $W_k(n_k) = \delta_k(\mu_k + \theta_k') - \delta_k'\theta_k' + W_k^{HT}(n_k) + o(1)$ as $\lambda_k \to \infty$.*

3.4 Multi-class M/M/1 queue

The multi-class M/M/1 queue has received lot of attention from the research community. In the case of linear holding cost, the $c\mu$-index rule has been proved to be optimal in two main settings: (i) with exponential distributed service times and preemptive scheduling Buyukkoc *et al.* [33], and (ii) general service time distributions and non-preemptive scheduling Gelenbe *et al.* [46]. A brief explanation of the optimality of an index rule is that having a linear holding cost c_k for a class-k customer per unit of time is equivalent to a problem where a reward c_k is received upon service completion (and no holding cost) Gittins *et al.* [48, Section 4.9]. The latter can be seen as a MABP, for which an index rule (in this case $c\mu$) is optimal[1]. However, this equivalence holds only for linear holding costs, which explains why for general cost functions the structure of the optimal scheduling policy is no longer of index type. In that context, a fruitful approach has been to derive scheduling policies with near-optimal performance or asymptotically optimal performance in a limiting regime, see the references as stated in Chapter 1.

In this section, we derive an index policy for the multi-class M/M/1 system by considering the limit of our Whittle index as the abandonment rate tends to 0. Note that Whittle's index $W_k(n_k)$ goes to ∞ as $\theta_k \to 0$, and it turns out that when scaling the index by θ_k we get a non-trivial limit. The proof of the next proposition may be found in Appendix 3.7.6.

Proposition 3.7. *Assume $\sum_{k=1}^K \rho_k < 1$, consequently $\rho_k < 1$. Assume $C_k(n_k, a) = C_k(n_k)$, $a = 0, 1$, $\theta_k' = \theta_k$, and $\delta_k = \delta_k' = 0$. Then,*

$$\lim_{\theta_k \to 0} \theta_k W_k(n_k) = \frac{\mu_k(1-\rho_k)}{\rho_k} \cdot \left[\sum_{m=0}^{\infty} \rho_k^m (1-\rho_k) C_k(n_k - 1 + m) - C_k(n_k - 1)\right]. \quad (3.4.1)$$

Observe that convexity of the function $C(\cdot)$ implies that (3.4.1) is a non-decreasing function.

A heuristic for the M/M/1 queue with as objective to minimize the holding cost can now be derived as follows. Set $\theta_k = \theta_k'$ for all k and consider the index multiplied by θ_k as $\theta_k \to 0$. A heuristic is then to give priority according to the index as given in (3.4.1).

[1] This is known as the *tax* formulation of a MABP, see [48, Section 4.9].

Note that it is not possible to derive the index policy for the standard M/M/1 queue in a direct fashion, which we will explain now. Consider the M/M/1 queue under threshold policy n with 0-1 structure. This system is equivalent to the FIFO M/M/1 queue with n permanent customers. It is known that the probability that the stationary process is in state n is then $1-\rho_k$. Hence, $\sum_{m=0}^{n}\pi_k^n(m) = 1-\rho_k$ for all n. The subsidy obtained on average is therefore $W(1-\rho_k)$, which is independent of the threshold policy n. This implies that the subsidy does not allow us to discriminate between different thresholds. We overcome this issue in Proposition 3.7 by considering abandonments to occur in the queue, and computing Whittle's index as $\theta_k \to 0$. In Gittins *et al.* [48, Section 6.5] the authors obtained the same index policy. They first derived Whittle's index for the total discounted holding cost criterion. In that case too, indexability needs to be established. Then taking the discount factor to zero resulted in the index policy as stated in Proposition 3.7.

For large values of n_k, the index (3.4.1) is approximately equal to $C_k'(n_k)\mu_k$, which we refer to as the *Gcμ*-rule. This rule was introduced in van Mieghem [70] for convex delay cost. The equivalence with the *Gcμ* rule can be seen as follows. We have for n_k large,

$$\sum_{m=0}^{\infty}\rho_k^m(1-\rho_k)C_k(n_k-1+m) - C_k(n_k-1)\sum_{m=0}^{\infty}\rho_k^m(1-\rho_k)$$
$$= (1-\rho_k)\sum_{m=0}^{\infty}\rho_k^m(C(n_k-1+m)-C(n_k-1)) \approx (1-\rho_k)\sum_{m=0}^{\infty} m\rho_k^m C'(n_k-1) \approx C'(n_k)\frac{\rho_k}{(1-\rho_k)},$$

where we used that for n_k large with respect to m, we have $\frac{C(n_k-1+m)-C(n_k-1)}{m} \approx C'(n_k)$ and that large values of m have a negligible weight on the summation. Hence, it follows from (3.4.1) that $\lim_{\theta_k\to 0}\theta_k W_k(n_k) \approx C_k'(n_k)\mu_k$.

Numerical example. In Table 3.2 we compare the suboptimality of the $C'(n)\mu$-rule and index-rule (3.4.1) in an M/M/1 queue without abandonments. Note that when $\theta_k = 0$, for all k, we need to assume $\sum_{k=1}^{K}\rho_k < 1$ in order to assure stability of the system. Consider 4 classes of customers with the following parameters: $\mu_1 = 16, \mu_2 = 27, \mu_3 = 12$ and $\mu_4 = 21$, $\rho_1 = 3\rho/9, \rho_2 = \rho/9$, $\rho_3 = 5\rho/9$ and $\rho_4 = \rho/9$. The holding cost of each class are cubic, $C_k(n_k) := \alpha_k + \beta_k n_k + \gamma_k n_k^2 + \delta_k n_k^3$, for which (3.4.1) simplifies to:

$$\beta_k\mu_k + \gamma_k\mu_k\left(\frac{3\rho_k-1}{1-\rho_k}+2n_k\right) + \delta_k\mu_k\left(3n_k^2+3\left(\frac{2\rho_k-1}{1-\rho_k}\right)n_k + \frac{4\rho_k^2+\rho_k+1}{(1-\rho_k)^2}\right).$$

We take the particular example: $C_1(n_1) = 6n_1+2n_1^2+2n_1^3$, $C_2(n_2) = 2n_2+2n_2^2+2n_2^3$, $C_3(n_3) = n_3+n_3^2+3n_3^3$ and $C_4(n_4) = 8n_4+2n_4^3$. We observe that for this example the $C'(n)\mu$-rule is outperformed by the index-rule (3.4.1), but both policies give nearly optimal performance for low workloads.

ρ	0.11	0.21	0.31	0.41
(3.4.1)	4.25e-06	1.51e-05	6.07e-06	5.02e-07
$C'(n)\mu$	0.0072	0.0636	0.1002	0.1320
ρ	0.51	0.61	0.71	0.81
(3.4.1)	0.008	0.0291	0.0919	1.7129
$C'(n)\mu$	0.1689	0.3616	1.8280	4.9539

Table 3.2: Absolute error $|\mathcal{C}^{\phi} - \mathcal{C}^{OPT}|$, with $\phi \in \{\phi_{(3.4.1)}, C'(n)\mu\}$, where $\mathcal{C}^{\phi_{(3.4.1)}}$ and $\mathcal{C}^{C'(n)\mu}$ are the average costs under the (3.4.1) policy and the $Gc\mu$-rule, respectively.

3.5 Asymptotic optimality

In this section we will discuss the asymptotic optimality of Whittle's index policy for the model under study. Section 3.5.1 deals with the optimality of Whittle's index policy in a multi-server setting, and Section 3.5.2 provides a method to prove Whittle's index policy to be optimal in light-traffic and heavy-traffic regimes.

3.5.1 Multi-server setting

For linear holding cost, asymptotic optimality in a multi-server setting can be directly derived from Verloop [93]. Assume there are M servers and the arrival rate of class-k customers is $M\lambda_k$. Let W_k be the state-independent index as given in (3.3.1). In Verloop [93, Proposition 6.2] it is shown that the Whittle index policy (WI), where at each moment in time a server serves a customers having highest non-negative index W_k, is asymptotically optimal in the following sense: for any policy ϕ,

$$\lim_{M\to\infty} \mathcal{C}^{WI}(M) \le \liminf_{M\to\infty} \overline{\mathcal{C}}^{\phi}(M),$$

where $\mathcal{C}^{WI}(M)$ denotes the average cost incurred by Whittle's index, and $\overline{\mathcal{C}}^{\phi}(M)$ denotes the average cost incurred by policy ϕ when there are M servers in the system.

For general holding cost, we can not derive asymptotic optimality. We do expect however that under certain conditions one would have the following. Assume there are M servers and x_kM queues where class-k customers arrive with rate λ_k, $k = 1, \ldots, K$[2]. A queue can be served by at most one server. In bandit terminology this represents having x_kM class-k bandits whose state (that is, the number of customers in the queue) has values in $E = \{0, 1, \ldots\}$, and the scheduler needs to decide which M bandits to make active (so which M queues to serve). In case the state space E would have been finite, the result in Verloop [93], Weber *et al.* [96] implies (under certain conditions) asymptotic optimality of Whittle's index policy as $M \to \infty$. However, for infinite state space, as is the case for our model, no result is known so far.

3.5.2 Light-traffic and heavy-traffic regimes

Light traffic and heavy traffic refer to the situations in which the total arrival rate goes to 0 and ∞, respectively. In order to take the limits we will modify the total arrival rate while keeping constant the

[2]This can represent for example a setting where there are x_kM class-k flows having newly arriving packets (represented by customers).

proportion of traffic of each class. To do so, we assume that $\lambda_k = \gamma_k \lambda$, where λ denotes the total arrival rate, and $\sum_{k=1}^{K} \gamma_k = 1$.

In this section we use the method introduced in Section 2.5 in order to show Whittle's index to be asymptotically optimal in both light-traffic and heavy-traffic regimes. In light traffic, most of the time the system is empty or at most there is one customer in the system. This implies that as $\lambda \to 0$, $REL(0)$ becomes admissible for the original problem, that is, $REL(0) \in \mathcal{U}$. In heavy traffic, we will prove that for the correct choice of W, under the Whittle index policy, constraint (2.3.1) is satisfied with equality, and $REL(W) \in \mathcal{U}$.

We present the asymptotic optimality result in the light-traffic regime in Proposition 3.8 and in Proposition 3.9 that corresponding to the heavy-traffic regime for the particular case of linear holding cost. The proofs can be found in Appendix 3.7.7 and Appendix 3.7.8, respectively.

Proposition 3.8. *Assume $C_k(0,0) \geq C_k(0,1)$, for all k. The Whittle index policy (WI) is asymptotically relatively optimal in light traffic, that is,*

$$\lim_{\lambda \downarrow 0} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} = 0,$$

with $\lambda_k = \lambda \gamma_k$, $\sum_{k=1}^{K} \gamma_k = 1$.

Proposition 3.9. *Assume holding costs to be linear, that is, $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k \min(n_k, a)$. Whittle's index $W_k(n)$ is then given by Proposition 3.3. Assume there exists $\bar{k} \in \{1, \ldots, K\}$ such that $W_{\bar{k}}(n) > W_k(n)$ for all $k \neq \bar{k}$. Then, the Whittle index policy (WI) is asymptotically optimal in heavy traffic, that is,*

$$\lim_{\lambda \to \infty} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} = 0,$$

with $\lambda_k = \lambda \gamma_k$, $\sum_{k=1}^{K} \gamma_k = 1$.

We see from the proof of Proposition 3.9 that in fact any policy that gives strict priority to class $\bar{k}$ will be optimal as $\lambda \uparrow \infty$.

3.6 Numerical results

The objective of the present section is to show in which regimes the Whittle index policy $W(n)$ (Equation (2.3.6)) performs well. We will focus on holding cost functions of the shape $C_k(n_k, a) = C_k(n_k)$ or $C_k(n_k, a) = C_k((n_k - a)^+)$, that is, the holding cost is a function of the number of class-k customers in the *system* or *queue* respectively. Hence, $\tilde{C}_k(n_k, a)$ reduces to $C_k(n_k) + \delta_k \theta_k n_k$ or $C_k((n_k - a)^+) + \delta_k \theta_k (n_k - a)^+ + \delta'_k \theta'_k \min(a, n_k)$, respectively.

In Section 3.6.1 we compare the structure of Whittle's index policy with the structure of the optimal policy, numerically. In Section 3.6.2 we then numerically compare the performance of the index policies with that of the optimal policy.

3.6.1 Structure of different policies

We compare the structure of the different index policies and the optimal policy for linear and convex holding cost.

Linear holding cost

By value iteration, see Section 1.3.3, we observed that for a wide range of parameters the optimal policy, under linear holding cost, is of the following structure: when $(N_1, \ldots, N_K)$ is close enough to the origin (and N_i denotes the number of class-i customers in the system), it is optimal to prioritize classes according to the $\tilde{c}\mu$-rule, otherwise to prioritize classes according to the $\tilde{c}\mu/\theta$-rule, where $\tilde{c}_k := c_k + \delta_k\theta_k$, see Figure 3.1 (left) with $\epsilon = 0$ as described in the next section. Hence, the Whittle's index (which corresponds to the $\tilde{c}\mu/\theta$-rule in the linear case) captures the optimal action for states that are not too close to the origin.

General holding cost

To discuss the structure of index policies for general holding cost, we focus on two classes of customers ($K = 2$). In a state (N_1, N_2), the action taken by Whittle's index rule is to serve the class having highest value $W_k(N_k)$. Since $W_k(N_k)$ is an non-decreasing function, this implies that there is an increasing switching curve (SC) such that when (N_1, N_2) is below the SC, Whittle's index policy serves class 1 and for any state (N_1, N_2) above the curve the policy serves class 2. Note that for linear holding cost this switching curve collapses either to the vertical or horizontal axis.

By value iteration we observed that an optimal policy is as well of switching curve type. For example, in Figure 3.1 (left) we plot the switching curve of the optimal policy with the following holding cost: $C_1(n) = n + \epsilon n^2$ and $C_2(n) = n$ (parameters $\theta = \theta'$ and $\lambda = [9, 10], \mu = [14, 16], \theta = [2, 0.05], \delta = [4, 0.3]$). When $\epsilon = 0$, we obtain a decreasing switching curve, which describes the behavior of the optimal policy for linear cost as explained in Section 3.6.1. As ϵ becomes positive, the switching curve becomes increasing. In addition, ϵ becomes larger, and hence the quadratic cost of class 1 increases, and therefore, class 1 gets priority in a larger region.

We now compare the actions taken under Whittle's index policy and the optimal policy. We consider an example with quadratic costs $C_1(n) = (c_{11} + \delta_1\theta_1)n + c_{21}n^2$ and $C_2(n) = (c_{12} + \delta_2\theta_2)n + c_{22}n^2$, and set the following parameters $\theta = \theta'$ and $\mu = [15, 18]; \theta = [4, 7]; c_1 = [1, 4]; c_2 = [2, 1]; \delta = [8, 6.5]$. In Figures 3.1 (middle and right) we plot the optimal actions (obtained by value iteration) for load $\rho = 0.8$ and $\rho = 2.5$, respectively, and compare it to the actions taken under Whittle's index policy. We observe that the optimal policy can be described by a switching curve. In addition the optimal policy coincides with that of Whittle's index $W(n)$ in almost all the state space as the workload increases. We also plot the switching curve corresponding to the fluid index $w(n)$ and observe a very good fit. The fluid index $w(n)$, as well as the fluid index policy will be introduced in Chapter 4.

3.6.2 Performance evaluation

In this section we evaluate numerically the performance of the index policies. This is carried out by computing the relative suboptimality gap between the average cost of the optimal solution and an index policy. In order to compute this we use the value iteration algorithm.

We saw in Section 3.4 that the index policy with index (3.4.1) performs very well in an M/M/1 multi-class systems (when there are no abandonments). We considered cubic costs and 4 classes of customers and compared the Generalized index rule ($Gc\mu$) and the index-rule of (3.4.1) and we observed there that the latter performs slightly better than the $Gc\mu$-rule.

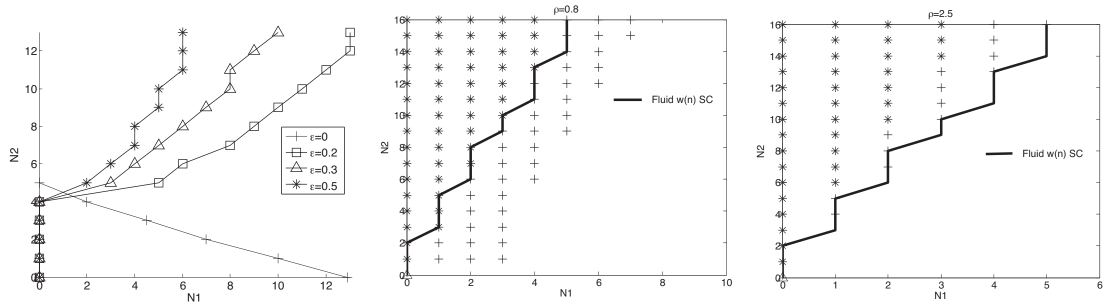


Figure 3.1: Left: Switching curves of the optimal policy for varying holding cost (from linear to quadratic). Middle and right: Actions under the optimal policy, the index policy $W(n)$, and the fluid index policy for quadratic holding cost. Area with "+": $W(n)$ serves class 1 while it is optimal to serve class 2, Area with "*": $W(n)$ serves class 2, which is also optimal, and in the white area $W(n)$ serve class 1, which is also optimal.

In this section we will consider scenarios allowing abandonments. We will evaluate the following indices: (i) the Whittle index $W(n)$ (Equation (2.3.6)), (ii) the Whittle index for large states $W^\infty(n)$ as proposed in Proposition 3.4, and (iii) the fluid index $w(n)$, which will be derived in Chapter 4. We compare these to the two index policies proposed for a multi-class queue without abandonments: the $Gc\mu$-rule, and the index-rule corresponding to (3.4.1) which is an approximation of $W(n)$ for θ close to zero. Note that both the $Gc\mu$-rule and the index-rule corresponding to (3.4.1) have been derived for the underload scenario ($\sum_{k=1}^K \rho_k < 1$) and are θ-independent. We do therefore not expect these index rules to perform near-optimal for high workloads.

We will analyze two different scenarios: (1) varying the workload ρ, and (2) varying the abandonment rates θ.

Varying Workload

In this section we aim at observing the behavior of index policies for varying workload. In order to do so we will compute the (relative) suboptimality gap of the proposed index policies. By suboptimality gap we mean $|\mathcal{C}^\phi - \mathcal{C}^{OPT}|$ and by relative suboptimality gap $(|\mathcal{C}^\phi - \mathcal{C}^{OPT}|)/\mathcal{C}^{OPT}$ where ϕ denotes the index policy under study.

Example with linear holding cost ($\theta = \theta'$)**:** We set $C_k(n,a) = c_k n$, $\mu = [15, 25]$, $\theta' = \theta = [4, 2]$, $c = [1, 1]$, $\delta = [5, 3.2]$, and let $\rho = \sum_{k=1}^2 \lambda_k/\mu_k$ vary in the interval $[0, 2.6]$, with $\lambda_1/\mu_1 = \lambda_2/\mu_2$. For linear holding costs, the indices $W(n), W^\infty(n)$ and $w(n)$ reduce to the $\tilde{c}\mu/\theta$-rule and the indices $Gc\mu$ and (3.4.1) reduce to the $\tilde{c}\mu$-rule, with $\tilde{c}_k = c_k + \delta_k\theta_k$.

Example with linear holding cost ($\theta \neq \theta'$)**:** We set $C_k(n,a) = c_k(n-a)^+$, $\mu = [15, 25]$, $\theta = [4, 2]$, $\theta' = [3, 2]$, $c' = c = [1, 1]$, $\delta = [5, 3.2]$, $\delta' = [2, 1]$, and let $\rho = \sum_{k=1}^2 \lambda_k/\mu_k$ vary in the interval $[0, 2.6]$, with $2\lambda_1/\mu_1 = \lambda_2/\mu_2$. For linear holding costs and $\theta \neq \theta'$, the indices $W(n), W^\infty(n)$ and $w(n)$ reduce to the $\tilde{c}(\mu + \theta')/\theta - \tilde{c}'$-rule and the indices $Gc\mu$ and (3.4.1) reduce to the $\tilde{c}\mu$-rule, with $\tilde{c}_k = c_k + \delta_k\theta_k$.

In Figure 3.2 we observe for both cases that the $\tilde{c}\mu$-rule is optimal in underload, while the performance of the index $W(n)$ is nearly optimal in overload, as expected from Proposition 3.9. As discussed in Section 3.6.1, in a state far from the origin, the optimal action is to serve according to $\tilde{c}\mu/\theta$, which is the region in which the process will live in overload, explaining why the $\tilde{c}\mu/\theta$-rule and the $\tilde{c}(\mu+\theta')/\theta - \tilde{c}'$-rule

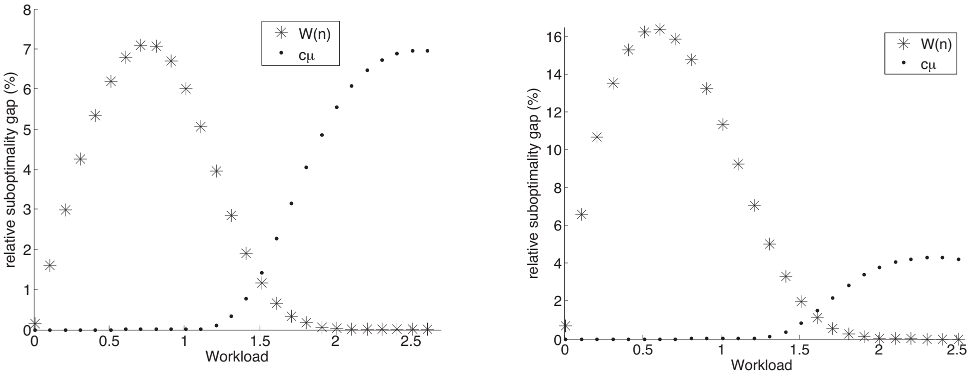


Figure 3.2: Left: linear holding cost, as ρ increases when $\theta = \theta'$. Right: linear holding cost, as ρ increases when $\theta \neq \theta'$.

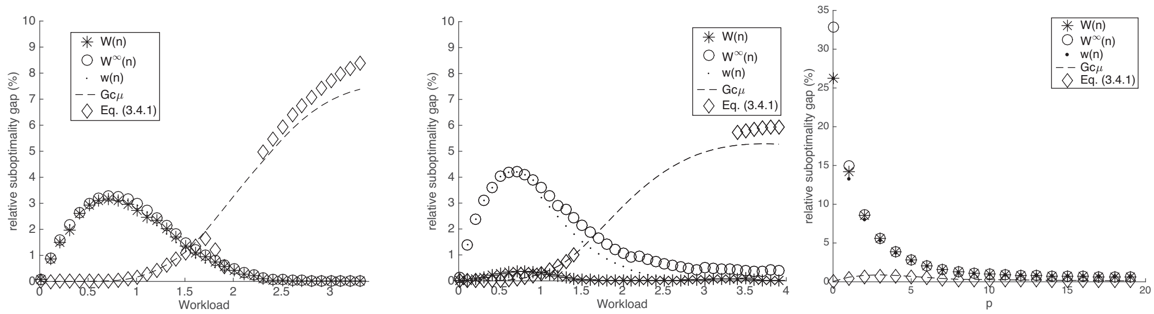


Figure 3.3: Left: linear holding cost, as ρ increases. Middle: quadratic holding cost as ρ increases. Right: quadratic holding cost as p ($\theta_i = p\epsilon_i, i \in \{1, 2\}$) varies.

perform well in this case. In underload, the effect of abandonments is not that important and the $\tilde{c}\mu$-rule performs very well.

Example with quadratic holding cost ($\theta = \theta'$)**:** Consider the following parameters: $\mu = [15, 18]$, $\theta' = \theta = [4, 7]$, $c_1 = [1, 4]$, $c_2 = [2, 1]$, $\delta = [8, 6.5]$, and we let λ vary, but keeping $\lambda_1/\mu_1 = \lambda_2/\mu_2$. We assume quadratic costs $C_1(n) = (c_{11} + \delta_1\theta_1)n + c_{21}n^2$ and $C_2(n) = (c_{12} + \delta_2\theta_2)n + c_{22}n^2$. See Figure 3.3 (left) for the sub-optimality gap and Table 3.3 for the absolute errors.

We observe that for low load the $Gc\mu$-rule and the index-rule (3.4.1) behave very well. However, as the load grows larger, the suboptimality gap of these θ-independent policies grows large, while our Whittle index policy $W(n)$, the Whittle index policy for large states $W^\infty(n)$ and the fluid index policy $w(n)$ become near optimal. In Table 3.3 we observe that the convergence towards optimality is reached very

Workload	1	1.5	2	2.5	3	3.5	5.25
$W(n)$	1.3089	1.4608	0.8055	0.1094	0.0185	0.0065	0.00017
$W^\infty(n)$	1.4028	1.5596	0.8902	0.1732	0.0614	0.0329	0.0007
$w(n)$	1.3823	1.2885	0.5534	0.0026	0.0771	0.0904	0.0004
(3.4.1)	0.0409	0.7327	0.8010	11.2134	20.5851	28.3926	50.0996
$Gc\mu$	0.0409	0.7483	3.9951	10.4111	18.7237	25.0454	42.5645

Table 3.3: Absolute error $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ that corresponds to the example in Figure 3.3 (left).

Workload	1	1.5	2.5	3	3.5	5.25	7.25	10	16
$W(n)$	0.1332	0.0664	0.0098	0.1260	0.2874	0.2448	0.1404	0.0486	0.0061
$W^{\infty}(n)$	1.4817	1.9167	1.4429	1.1485	1.4243	1.7296	1.4784	0.7977	0.1012
$w(n)$	1.4817	1.4157	0.3397	0.0382	0.1288	0.5125	04383	0.1542	0.0093
(3.4.1)	0.0720	-	-	-	19.3226	35.5180	48.5766	66.1024	91.4859
$Gc\mu$	0.0720	0.7896	7.7697	12.8528	17.6942	31.1417	43.3748	59.7161	99.4344

Table 3.4: Absolute error $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ that corresponds to the example in Figure 3.3 (middle).

fast as the absolute error ($\mathcal{C}^{WI} - \mathcal{C}^{OPT}$) of the $W(n), W^{\infty}(n)$ and $w(n)$ indices is of order 10^{-4} when $\rho = 5.25$. On the other hand, both (3.4.1) and the $Gc\mu$-rule perform very bad in overload. Hence, our index policies are very suitable for the overload setting, which are from a practical point of view of main importance.

Note that the jump around $\rho = 2$ for the index-rule (3.4.1) is a result of undefined values around $\lambda_k = \mu_k$.

Example with quadratic holding cost ($\theta \neq \theta'$)**:** Consider the following parameters: $\mu = [15, 18]$, $\theta = [4, 7]$, $\theta' = [3, 4]$, $c_1 = [1, 4]$, $c_2 = [2, 1]$, $\delta = [8, 6.5]$, $\delta' = [7, 7]$ and we let λ vary, but keeping $2\lambda_1/\mu_1 = \lambda_2/\mu_2$. We assume quadratic costs $\tilde{C}_1(n,a) = c_{11}(n-a)^+ + c_{21}((n-a)^+)^2 + \delta_1\theta_1(n-a)^+ + \delta_1'\theta_1'a$ and $\tilde{C}_2(n,a) = c_{12}(n-a)^+ + c_{22}((n-a)^+)^2 + \delta_2\theta_2(n-a)^+ + \delta_2'\theta_2 a$. See Figure 3.2 for the suboptimality gap and Table 3.4 for the absolute errors.

We observe that for low loads the $Gc\mu$-rule and the index-rule (3.4.1) behave very well. In this example, also the Whittle index policy performs close to optimal for low loads, while $W^{\infty}(n)$ and $w(n)$ do not. As the load grows larger, Whittle's index policy $W(n)$, and the fluid index policy $w(n)$ become near optimal. However, in this example the convergence towards optimality in absolute terms is much slower than for the previous example. The absolute error $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ is of order 10^{-3} for the indices $W(n)$ and $w(n)$ and of order 10^{-1} for $W^{\infty}(n)$ when $\rho = 16$. This phenomena is explained by the fact that the process lives around an area where the optimal policy prescribes to serve class-2 customers and the index policies prescribe to serve class-1 customers. As the workload increases this phenomena disappears.

The jump around the interval $\rho = [1.5, 3]$ for the index-rule (3.4.1) is a result of undefined values around $\lambda_k = \mu_k$.

Varying abandonment rates

In this section we evaluate the performance of the index policies for varying abandonment rates.

Linear holding cost: In this case, the five index policies mentioned above reduce to the $\tilde{c}\mu/\theta$-rule and the $\tilde{c}\mu$-rule, as explained in Section 3.6.2. As $\theta_k \to 0$, we observed in the numerical experiments that the $\tilde{c}\mu$-rule performs optimal, while the $\tilde{c}\mu/\theta$-rule might perform very bad when the abandonment rates are negligibly small. It is known that for the non-reneging case, the $\tilde{c}\mu$-rule is optimal in underload (the celebrated $c\mu$-rule for a multi-class M/M/1 queue). The $\tilde{c}\mu/\theta = (c + \delta\theta)\mu/\theta$ index might however give an opposite priority rule when θ's are very small, which explains the non-optimality of the $\tilde{c}\mu/\theta$-rule when θ_k's are very small.

Quadratic holding cost: Consider a system with two classes of customers. We assume quadratic holding costs $C_1(n) = \tilde{c}_{11}n + c_{21}n^2$ where, $\tilde{c}_{11} = (c_{11} + \delta_1\theta_1)$, and $C_2(n) = \tilde{c}_{21}n + c_{22}n^2$, where $\tilde{c}_{21} =$

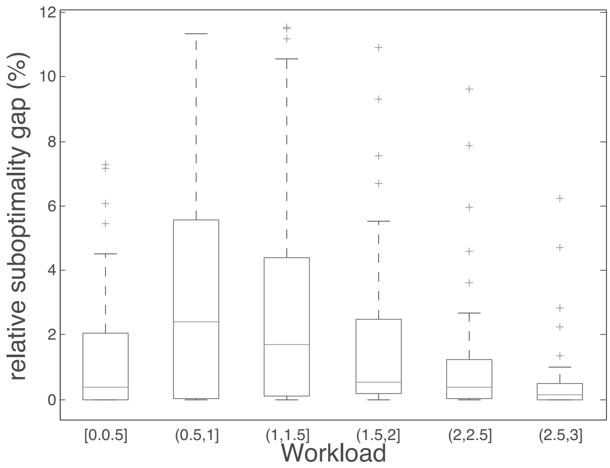


Figure 3.4: Suboptimality gap of Whittles index policy, for randomly generated parameters. The edges of the box represent the 25th and 75th percentile, the line inside the box the mean value corresponding to all values in that box and the "+"s are the outliers.

$(c_{21} + \delta_2\theta_2)$ and fix the following parameters: $\lambda = [4, 5]$, $\mu = [15, 17]$, $c_1 = [1, 4]$, $c_2 = [5, 1]$, $\delta = [2, 3]$, $\theta_1 = \epsilon_1 p$ and $\theta_2 = \epsilon_2 p$, where $\epsilon_1 = 0.05$ and $\epsilon_2 = 0.01$, and let p vary.

In Figures 3.3 (right) we plot the suboptimality gap as p varies from 0 to 200, hence θ_1 and θ_2 range from $[0, 10]$ and $[0, 2]$, respectively. We observe for the θ-dependent indices a suboptimality gap of 25% around $p = 0$. As θ grows large, this gap disappears however very fast. Note that the θ-independent indices work well, as we are in an underload scenario.

Example with random samples

We consider 360 samples with random values of $\lambda_k, \mu_k, \theta'_k, \theta_k$ and $c_k = [c_{k1}, c_{k2}]$ for $k = 1, 2$, where c_k is defined in the previous example. We compute the relative sub-optimality gap of Whittle's index policy, see Figure 3.4. Figure 3.4 is a box plot, where each box contains the suboptimality gap of 60 samples. We group the results in workload intervals of length 0.5. We observe that the average performance of Whittle's index policy is nearly optimal for high workloads, whereas the sub-optimality gap is largest for values of the workload in $(0.5, 1]$.

3.7 Appendix

3.7.1 Proof of Proposition 3.1

In Proposition 3.1 we aim at proving that threshold policy $\phi = n$ is an optimal solution of problem (2.3.3). In order to do so, we are left to prove the convexity of the value function V. We will therefore prove that the value function that corresponds to the truncated system $V^L(m)$ (truncated by $L > 1$) is convex. Having done this, due to the result in Bhulai *et al.* [25, Th. 3.1] we have that $V^L(m) \to V(m)$ as $L \to \infty$ and hence, the convexity of V^L for all L will imply convexity of the function V. In order to apply Bhulai *et al.* [25, Th. 3.1] we need to make sure that the conditions required are satisfied. Therefore, we first check the conditions required by Bhulai *et al.* [25, Th. 3.1], and then we prove the convexity of V^L.

Conditions to be checked for Bhulai *et al.* [25, Th. 3.1]

Let us first present the following definition:

Definition 3.1. *A function $f : E \longrightarrow \mathbb{R}_+$ is a moment function if there exists an increasing sequence of finite sets $E_r \uparrow E$, $r \to \infty$, such that* $\inf\{f(m) : m \notin E_r\} \to \infty$ *as $r \to \infty$. (Where E is the state space).*

Let us define $q^{\phi,L}(m, m-1) = \mu S^\phi(m) + \theta' S^\phi(m) - \theta(m - S^\phi(m))$, and recall that $q^{\phi,L}(m, m+1) = \lambda\left(1 - \frac{m}{L}\right)^+$. The conditions to be checked in Bhulai *et al.* [25, Th. 3.1] are the following:

1. There exists a moment function $f : \mathbb{N} \cup \{0\} \longrightarrow \mathbb{R}_+$, constants $\alpha, \beta > 0$ and $\tilde{M} > 0$ such that

$$\sum_{\tilde{m}=0}^{\infty} q^{\phi,L}(m, \tilde{m}) f(\tilde{m}) \leq -\alpha f(m) + \beta \mathbf{1}_{\{m<\tilde{M}\}}(m), \text{ for all } \phi, L \text{ and all } m,$$

   where ϕ defines the policy followed, L is the truncating parameter and $q^{\phi,L}(m, \tilde{m})$ the transition rate from m to $\tilde{m}$ under ϕ and L.

2. $(S^\phi(m), L) \mapsto q^{\phi,L}(m, \tilde{m})$ and $(S^\phi(m), L) \mapsto \sum_{\tilde{m}} q^{\phi,L}(m, \tilde{m}) f(\tilde{m})$ are continuous functions in $S^\phi(m)$ and L for all m and $\tilde{m}$.

We define $f(m) := \mathrm{e}^{\epsilon m}$, where $\epsilon > 0$. We can construct $E_r = \{0, \dots, r\}$ such that E_r is finite, $E_r \uparrow \mathbb{N} \cup \{0\}$ as $r \to \infty$ and $\inf\{f(m) : m \notin E_r\} \to \infty$. The objective is then to see, that there exists $\epsilon > 0$, an $\tilde{M} > 0$ and a constant $\alpha > 0$, such that

$$\sum_{\tilde{m}=0}^{\infty} q^{\phi,L}(m, \tilde{m}) f(\tilde{m}) \leq -\alpha f(m), \text{ for all } m \geq \tilde{M},$$

that is,

$$\begin{aligned}&\lambda\left(1 - \frac{m}{L}\right) \mathrm{e}^{\epsilon(m+1)} + \big((\mu + \theta')S^\phi(m) + \theta(m - S^\phi(m))\big)\, \mathrm{e}^{\epsilon(m-1)} \\ &- \left(\left(1 - \frac{m}{L}\right) + (\mu + \theta')S^\phi(m) + \theta(m - S^\phi(m))\right) \mathrm{e}^{\epsilon m} \leq -\alpha \mathrm{e}^{\epsilon m}, \text{ for all } m \geq \tilde{M}.\end{aligned}$$

After some algebra we get

$$\lambda\left(1 - \frac{m}{L}\right)(\mathrm{e}^{\epsilon} - 1) + \big((\mu + \theta' - \theta)S^\phi(m) + \theta m\big)(\mathrm{e}^{-\epsilon} - 1) \leq -\alpha, \text{ for all } m \geq \tilde{M}.$$

Note that $\lambda(1-m/L)(\mathrm{e}^\epsilon - 1)$ can be upper bounded by a constant, $\kappa_1 = \lambda(\mathrm{e}^\epsilon - 1)$, and $(\mu + \theta' - \theta)S^\phi(m)(\mathrm{e}^\epsilon - 1) \leq 0$. Besides, $\theta m(\mathrm{e}^{-\epsilon} - 1) < 0$. Hence, we can find $\tilde{M}$ large enough so that $-\theta m(\mathrm{e}^{-\epsilon} - 1) \geq \kappa_1$ for all $m \geq \tilde{M}$. This proves that condition (1) is satisfied.

Condition (2), i.e., the continuity of the functions $(S^\phi(m), L) \mapsto q^{\phi,L}(m, \tilde{m})$ and $(S^\phi(m), L) \mapsto \sum_{\tilde{m}} q^{\phi,L}(m, \tilde{m}) f(\tilde{m})$ in L and $S^\phi(m)$ is satisfied by definition of transition rates.

Convexity of V^L

For the ease of clarity we define $\omega := \mu+\theta'-\theta$ throughout this proof. W.l.o.g. assume $\lambda+\mu+\theta'+\theta L = 1$. For $n \in \{0,1,\dots,L\}$ we define $V_t^L(n)$ by $V_0^L(n) = 0$ and

$$\begin{aligned} V_{t+1}^L(n) =& \lambda\left(1-\frac{n}{L}\right)V_t^L(\min\{n+1,L\}) \\ &+ \min\{-W+\tilde{C}(n,0)+\omega V_t^L(n), \tilde{C}(n,1)+\omega V_t^L((n-1)^+)\} \\ &+ \theta n V_t^L((n-1)^+) + \lambda\frac{n}{L}V_t^L(n) + (L-n+1)\theta V_t^L(n). \end{aligned}$$

Where the right hand side is the dynamic programming operator to V_t^L. We will prove that V_t^L is a convex function for $n \le L-1$, that is,

$$2V_t^L(n) \le V_t^L((n-1)^+) + V_t^L(n+1), \text{ for } n \le L-1. \tag{3.7.1}$$

The function V_t^L being convex, for any t, will imply convexity of V^L and concludes the proof.

In order to prove convexity of V_t^L we first prove that $V_t^L(\cdot)$ is a non-decreasing function. The proof follows by induction: $V_0^L(n) = 0$ is non-decreasing for $t = 0$, then we assume $V_t^L(n)$ is non-decreasing and we prove that

$$V_{t+1}^L(n+1) - V_{t+1}^L(n) \ge 0 \text{ for all } n \le L-1. \tag{3.7.2}$$

Let us first consider the terms multiplied by λ in $V_{t+1}^L(n+1) - V_{t+1}^L(n)$, that is,

$$\begin{aligned} &\lambda\left(1-\frac{n+1}{L}\right)V_t^L(\min\{n+2,L\}) + \lambda\frac{n+1}{L}V_t^L(\min\{n+1,L\}) \\ &- \lambda\left(1-\frac{n}{L}\right)V_t^L(\min\{n+1,L\}) - \lambda\frac{n}{L}V_t^L(n) \\ &\ge \lambda\left(1-\frac{n+1}{L}\right)(V_t^L(\min\{n+2,L\}) - V_t^L(\min\{n+1,L\})) \\ &\quad + \lambda\frac{n}{L}(V_t^L(\min\{n+1,L\} - V_t^L(n))) \ge 0, \end{aligned}$$

where the last inequality holds due to the non-decreasingness of $V_t^L(n)$. Let us now consider the terms multiplied by θ in $V_{t+1}^L(n+1) - V_{t+1}^L(n)$, namely,

$$\begin{aligned} &\theta(n+1)V_t^L(n) + (L-n-1)\theta V_t^L(\min\{n+1,L\}) - \theta n V_t^L((n-1)^+) - (L-n)\theta V_t^L(n) \\ &\ge \theta n(V_t^L(n) - V_t^L((n-1)^+)) + (L-n-1)(V_t^L(\min\{n+1,L\}) - V_t^L(n)) \ge 0, \end{aligned}$$

where, again, the last inequality holds due to $V_t^L(n)$ being non-decreasing for all $n \le L-1$. Finally, let us consider the min-terms in $V_{t+1}^L(n+1) - V_{t+1}^L(n)$. It is straightforward that

$$\begin{aligned} &\min\{-W+\tilde{C}(\min\{n+1,L\},0) + (\mu+\theta')V_t^L(\min\{n+1,L\}), \\ &\qquad \tilde{C}(\min\{n+1,L\},1) + (\mu+\theta')V_t^L(n)\} \\ &- \min\{-W+\tilde{C}(n,0)+(\mu+\theta')V_t^L(n), \\ &\qquad \tilde{C}(n,1) + (\mu+\theta')V_t^L((n-1)^+)\} \ge 0, \end{aligned}$$

due to $\tilde{C}$ and V_t^L being non-decreasing. This proves (3.7.2) and hence we showed that $V_t^L(n)$ is non-decreasing.

Equation (3.7.1) for $n = 0$ follows directly from $V_t^L(\cdot)$ being non-decreasing. In the remainder of the proof we therefore prove Equation (3.7.1) for $n \geq 1$.

We will prove convexity (3.7.1) by induction on t. Since $V_0^L(n) = 0$, it holds for $t = 0$. Now assume V_t^L is convex. For $1 \leq n \leq L-1$ we have

$$\begin{aligned} 2V_{t+1}^L(n) =& 2\lambda\left(1-\frac{n}{L}\right)V_t^L(n+1) + 2\lambda\frac{n}{L}V_t^L(n) + 2\theta n V_t^L(n-1) + 2(L-n+1)\theta V_t^L(n) \\ &+ 2\min\{-W + \tilde{C}(n,0) + \omega V_t^L(n), \tilde{C}(n,1) + \omega V_t^L(n-1)\}. \end{aligned} \tag{3.7.3}$$

We need to show that this is less than or equal to $V_{t+1}^L(n-1) + V_{t+1}^L(n+1)$, which is given by

$$\begin{aligned} &\lambda\left(1-\frac{n-1}{L}\right)V_t^L(n) + \lambda\left(1-\frac{n+1}{L}\right)V_t^L(n+2) + \lambda\frac{n-1}{L}V_t^L(n-1) + \lambda\frac{n+1}{L}V_t^L(n+1) \\ &+ \theta(n-1)V_t^L((n-2)^+) + \theta(n+1)V_t^L(n) + (L-n+2)\theta V_t^L(n-1) + (L-n)\theta V_t^L(n+1) \\ &+ \min\{-W + \tilde{C}(n-1,0) + \omega V_t^L(n-1), \tilde{C}(n-1,1) + \omega V_t^L((n-2)^+)\} \\ &+ \min\{-W + \tilde{C}(n+1,0) + \omega V_t^L(n+1), \tilde{C}(n+1,1) + \omega V_t^L(n)\}. \end{aligned} \tag{3.7.4}$$

We first consider the two terms multiplied by λ in (3.7.3) and show that they are smaller than or equal to

$$\lambda\left(1-\frac{n-1}{L}\right)V_t^L(n) + \lambda\left(1-\frac{n+1}{L}\right)V_t^L(n+2) + \lambda\frac{n-1}{L}V_t^L(n-1) + \lambda\frac{n+1}{L}V_t^L(n+1). \tag{3.7.5}$$

When $1 \leq n < L-1$, then for the terms multiplied by λ in (3.7.3) we can write

$$\begin{aligned} 2\left(1-\frac{n}{L}\right)V_t^L(n+1) + 2\frac{n}{L}V_t^L(n) &= 2\left(1-\frac{n+1}{L}\right)V_t^L(n+1) + 2\frac{n}{L}V_t^L(n) + \frac{2}{L}V_t^L(n+1) \\ \leq \left(1-\frac{n-1}{L}\right)V_t^L(n) - \frac{2}{L}V_t^L(n) &+ \left(1-\frac{n+1}{L}\right)V_t^L(n+2) + 2\frac{n}{L}V_t^L(n) + \frac{2}{L}V_t^L(n+1), \end{aligned} \tag{3.7.6}$$

by convexity of V_t^L. Since by convexity $2\frac{n-1}{L}V_t^L(n) \leq \frac{n-1}{L}(V_t^L(n-1) + V_t^L(n+1))$, we obtain that (3.7.6) is smaller than or equal to (3.7.5). When $n = L-1$, it reduces to verifying $2(1-2/L)V_t^L(L-1) \leq (1-2/L)(V_t^L(L-2) + V_t^L(L))$, which follows from convexity of V_t^L.

For the terms multiplied by θ, we need to show for $1 \leq n \leq L-1$ that

$$\begin{aligned} &2nV_t^L(n-1) + 2V_t^L(n) + 2(L-n)V_t^L(n) \\ &\leq (n-1)V_t^L((n-2)^+) + (n+1)V_t^L(n) + 2V_t^L(n-1) + (L-n)(V_t^L(n-1) + V_t^L(n+1)). \end{aligned}$$

We apply the inequality $2V_t^L(n-1) \leq V_t^L((n-2)^+) + V_t^L(n)$ on the right hand side and the whole initial inequality reduces to

$$2nV_t^L(n-1) + 2(L-n)V_t^L(n) \leq n(V_t^L((n-2)^+) + V_t^L(n)) + (L-n)(V_t^L(n-1) + V_t^L(n+1)),$$

which holds by convexity of V_t^L.

We now consider the min-terms. We will condition on the possible optimal actions in states $n-1$ and $n+1$. Since at time t we have that V_t^L is convex, the optimal actions satisfy the monotonicity property. Denote by $a_n^* \in \{0,1\}$ the optimal action in state n, with action 0 (1) being passive (active). Then, by monotonicity there are the following three possibilities: (a_{n-1}^*, a_{n+1}^*) equals $(0,0), (0,1)$ or $(1,1)$. First assume $a^* = (0,1)$. Then, we obtain for $1 \leq n \leq L-1$ that

$$\begin{aligned}
&2\min\{-W + \tilde{C}(n,0) + \omega V_t^L(n), \tilde{C}(n,1) + \omega V_t^L(n-1)\} \\
&\leq -W + \tilde{C}(n,0) + \omega V_t^L(n) + \tilde{C}(n,1) + \omega V_t^L(n-1) \\
&\leq -W + \tilde{C}(n-1,0) + \omega V_t^L(n) + \tilde{C}(n+1,1) + \omega V_t^L(n-1) \\
&= \min\{-W + \tilde{C}(n-1,0) + \omega V_t^L(n-1), \tilde{C}(n-1,1) + \omega V_t^L((n-2)^+)\}, \\
&\quad + \min\{-W + \tilde{C}(n+1,0) + \omega V_t^L(n+1), \tilde{C}(n+1,1) + \omega V_t^L(n)\},
\end{aligned} \tag{3.7.7}$$

where in the second inequality we used that C and hence $\tilde{C}$ satisfies (3.1.2). In the case $a^* = (1,1)$ we obtain for $1 \leq n \leq L-1$ that

$$\begin{aligned}
&2\min\{-W + \tilde{C}(n,0) + \omega V_t^L(n), \tilde{C}(n,1) + \omega V_t^L(n-1)\} \\
&\leq 2\tilde{C}(n,1) + 2\omega V_t^L(n-1) \\
&\leq \tilde{C}(n-1,1) + \tilde{C}(n+1,1) + \omega(V_t^L((n-2)^+) + V_t^L(n)) \\
&= \min\{-W + \tilde{C}(n-1,0) + \omega V_t^L(n-1), \tilde{C}(n-1,1) + \omega V_t^L((n-2)^+)\}, \\
&\quad + \min\{-W + \tilde{C}(n+1,0) + \omega V_t^L(n+1), \tilde{C}(n+1,1) + \omega V_t^L(n)\}.
\end{aligned} \tag{3.7.8}$$

In the second inequality we used the convexity of C (and hence of $\tilde{C}$) and the convexity of V_t^L.

When $a^* = (0,0)$ we obtain for $1 \leq n \leq L-1$ that

$$\begin{aligned}
&2\min\{-W + \tilde{C}(n,0) + \omega V_t^L(n), \tilde{C}(n,1) + \omega V_t^L(n-1)\} \\
&\leq -2W + 2\tilde{C}(n,0) + 2\omega V_t^L(n) \\
&\leq -2W + \tilde{C}(n-1,0) + \tilde{C}(n+1,0) + \omega V_t^L(n-1) + \omega V_t^L(n+1) \\
&= \min\{-W + \tilde{C}(n-1,0) + \omega V_t^L(n-1), \tilde{C}(n-1,1) + \omega V_t^L((n-2)^+)\}, \\
&\quad + \min\{-W + \tilde{C}(n+1,0) + \omega V_t^L(n+1), \tilde{C}(n+1,1) + \omega V_t^L(n)\}.
\end{aligned} \tag{3.7.9}$$

In the second inequality we used the convexity of C (and hence of $\tilde{C}$) and the convexity of V_t^L.

Hence, we have that (3.7.3) is less than or equal to $V_{t+1}^L(n-1) + V_{t+1}^L(n+1)$, hence V_{t+1}^L is convex. This concludes the proof for convexity of $V_t^L(\cdot)$. Since $V_t^L \to V^L$ as $t \to \infty$ Puterman [77, Chap. 9.4], convexity of $V_t^L(\cdot)$ implies convexity of $V^L(\cdot)$.

3.7.2 Proof of Proposition 3.3

For ease of notation we omit subscript k from the notation in the proof. To calculate Whittle's index as in Theorem 2.2 we need to consider the threshold policies n and $n-1$ in which the server is active in states $m \geq n+1$ and $m \geq n$, respectively.

Let us consider the policy n first. Let $f^n(ab)$ and $f^n(ser)$ denote the fraction of customers that end up abandoning and being served, respectively. A rate conservation argument implies that all arriving users either abandon or are served, thus $\lambda = \lambda f^n(ab) + \lambda f^n(ser)$. Conditioning on the state, the rate of abandonment from the system can be written as $\sum_{m=0}^{\infty} \theta m \pi^n(m) + (\theta' - \theta) \sum_{m=n+1}^{\infty} \pi^n(m)$, and equating the rates we get the relation

$$\theta \mathbb{E}(N^n) + (\theta' - \theta) \sum_{m=n+1}^{\infty} \pi^n(m) = \lambda f^n(ab) = \lambda(1 - f^n(ser)). \tag{3.7.10}$$

Conditioning on the state, the rate of service is given by $\sum_{m=n+1}^{\infty} \mu \pi^n(m)$, and we get the relation

$$\lambda f(ser) = \mu \sum_{m=n+1}^{\infty} \pi^n(m),$$

and substituting in (3.7.10) we get

$$\mathbb{E}(N^n) = \frac{\lambda}{\theta} + \frac{\theta - \theta' - \mu}{\theta} \sum_{m=n+1}^{\infty} \pi^n(m),$$

where N^n denotes the stationary number of class-k customers in the system under the threshold policy n. We calculate now the average holding cost. Plugging the holding cost $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k a$ in the total cost relation (3.1.4) we get $\tilde{C}(n, a) = \tilde{c}n + a(\tilde{c}' - \tilde{c})$, where the constants $\tilde{c}$ and $\tilde{c}'$ are defined in the statement. The average cost then becomes

$$\begin{aligned} &\mathbb{E}(\tilde{C}(N^n, S^n(N^n))) = \tilde{c}\mathbb{E}(N^n) + (\tilde{c}' - \tilde{c}) \sum_{m=n+1}^{\infty} \pi^n(m) \\ &= \tilde{c}\frac{\lambda}{\theta} + \left(\frac{\tilde{c}(\theta - \theta' - \mu)}{\theta} + \tilde{c}' - \tilde{c}\right) \sum_{m=n+1}^{\infty} \pi^n(m) = \tilde{c}\frac{\lambda}{\theta} + \left(\tilde{c}' - \frac{\tilde{c}(\theta' + \mu)}{\theta}\right) \sum_{m=n+1}^{\infty} \pi^n(m). \end{aligned}$$

We substitute now all the terms in (2.3.6) to get

$$W(n) = \frac{\tilde{c}(\mu + \theta')}{\theta} - \tilde{c}', \tag{3.7.11}$$

which concludes the proof.

3.7.3 Proof of Proposition 3.4

For ease of notation we drop the dependency on k throughout the proof.

The index in the case $\mu + \theta' = \theta$ was obtained in (3.2.6), therefore we assume $\mu + \theta' > \theta$ throughout the proof. First of all recall that the steady-state probabilities $\pi^n(i)$ for policy n and state i are given by (3.2.3). To compute Whittle's index for large values of n, we need to compute $\pi^n(i) - \pi^{n-1}(i), \forall i \geq 0$.

Let us start by $i = 0$, that is,

$$\begin{aligned}\pi^n(0) - \pi^{n-1}(0) =& \frac{\left(\pi^{n-1}(0)\right)^{-1} - \left(\pi^n(0)\right)^{-1}}{\left(\pi^n(0)\pi^{n-1}(0)\right)^{-1}} \\ =& \left(\sum_{i=1}^{\infty}\prod_{m=1}^{i}\frac{q^{n-1}(m-1,m)}{q^{n-1}(m,m-1)} - \sum_{i=1}^{\infty}\prod_{m=1}^{i}\frac{q^{n}(m-1,m)}{q^{n}(m,m-1)}\right)\pi^n(0)\pi^{n-1}(0).\end{aligned}$$

The following observations on the transition rates will be used throughout the proof:

$$q^n(m,m-1) = q^{n-1}(m,m-1),\ \ \forall\, m \neq n, m \geq 1, \tag{3.7.12}$$

$$q^n(m-1,m) = q^{n-1}(m-1,m),\ \ \forall\, m \geq 1. \tag{3.7.13}$$

Taking these relations into account together with the fact that $q^n(n,n-1) - q^{n-1}(n,n-1) = \theta - \mu - \theta'$, we get after some calculations

$$\pi^n(0) - \pi^{n-1}(0) = \pi^n(0)\pi^{n-1}(0)\sum_{i=n}^{\infty}\prod_{\substack{m=1\\ m\neq n}}^{i}\frac{q^{n}(m-1,m)}{q^{n}(m,m-1)}\left(\frac{1}{q^{n-1}(n,n-1)} - \frac{1}{q^{n}(n,n-1)}\right)$$

$$= \pi^n(0)\pi^{n-1}(0)\frac{\theta-\mu-\theta'}{q^{n-1}(n,n-1)}\sum_{i=n}^{\infty}\prod_{m=1}^{i}\frac{q^{n}(m-1,m)}{q^{n}(m,m-1)}.$$

Since $q^n(m-1,m) = \lambda$ for all $m \geq 1$, $q^n(m,m-1) = \theta m$ for all $1 \leq m \leq n-1$ and $q^n(m,m-1) = \mu + \theta' + \theta(m-1)$ for all $m \geq n$, together with $\pi^n(0)$ given as in (3.2.3), we observe that

$$\frac{\pi^n(0)\pi^{n-1}(0)}{q^{n-1}(n,n-1)} \in \mathcal{O}\left(\frac{1}{n}\right) \quad \text{and} \quad \sum_{i=n}^{\infty}\prod_{m=1}^{i}\frac{q^{n}(m-1,m)}{q^{n}(m,m-1)} \in \mathcal{O}\left(\frac{1}{n!}\right).$$

We then get that

$$\pi^n(0) - \pi^{n-1}(0) \in \mathcal{O}\left(\frac{1}{nn!}\right). \tag{3.7.14}$$

We can now compute $\pi^n(i) - \pi^{n-1}(i)$, for all $0 < i \leq n-1$. Using (3.7.13), we obtain for $i \leq n-1$,

$$\pi^n(i) - \pi^{n-1}(i) = \prod_{m=1}^{i}\frac{q^{n}(m-1,m)}{q^{n}(m,m-1)}(\pi^n(0) - \pi^{n-1}(0)).$$

Due to (3.7.14) and since $q^n(m,m-1) = \theta m$, and $q^n(m-1,m) = \lambda$ for all $m \leq n-1$, we obtain for $i \leq n-1$

$$\pi^n(i) - \pi^{n-1}(i) = \frac{\lambda^i}{i!\theta^i}(\pi^n(0) - \pi^{n-1}(0)) \in \mathcal{O}\left(\frac{1}{nn!}\right), \tag{3.7.15}$$

For states $i \geq n$, with n sufficiently large, we have the following:

$$\pi^n(i) - \pi^{n-1}(i) = \prod_{m=1}^{i}\frac{q^{n}(m-1,m)}{q^{n}(m,m-1)}\pi^n(0) - \prod_{m=1}^{i}\frac{q^{n-1}(m-1,m)}{q^{n-1}(m,m-1)}\left(\pi^n(0) - \pi^n(0) + \pi^{n-1}(0)\right).$$

From observation (3.7.14), together with $\prod_{m=1}^{i} \dfrac{q^{n-1}(m-1,m)}{q^{n-1}(m,m-1)} \in \mathcal{O}\left(\frac{1}{i!}\right)$, we obtain

$$\pi^n(i) - \pi^{n-1}(i) = \mathcal{O}\left(\frac{1}{i!n!n}\right) + \prod_{m=1}^{i} \frac{q^n(m-1,m)}{q^n(m,m-1)}\pi^n(0) - \prod_{m=1}^{i} \frac{q^{n-1}(m-1,m)}{q^{n-1}(m,m-1)}\pi^n(0).$$

After some calculations and by observations (3.7.12) and (3.7.13) we obtain

$$\begin{aligned} \pi^n(i) - \pi^{n-1}(i) &= \left(\frac{1}{q^n(n,n-1)} - \frac{1}{q^{n-1}(n,n-1)}\right) \prod_{\substack{m=1 \\ m\neq n}}^{i} \frac{q^n(m-1,m)}{q^n(m,m-1)} + \mathcal{O}\left(\frac{1}{i!n!n}\right) \\ &= \frac{\mu+\theta'-\theta}{q^{n-1}(n,n-1)}\pi^n(i) + \mathcal{O}\left(\frac{1}{i!n!n}\right), \end{aligned} \tag{3.7.16}$$

for $i \geq n$. Recall from (3.3.3) that Whittle's index can be written as $\delta(\mu+\theta') - \delta'\theta' + W^c(n)$, where $W^c(n)$ corresponds to the holding costs only. $W^c(n)$ can be written as follows

$$W^c(n) = \frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = \frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n) + \mathcal{O}(1/n!n)}, \tag{3.7.17}$$

with

$$\begin{aligned} \xi_1(n) &:= \sum_{i=1}^{n-1} C(i,0)(\pi^n(i) - \pi^{n-1}(i)), \\ \xi_2(n) &:= C(n,0)\pi^n(n) - C(n,1)\pi^{n-1}(n), \\ \xi_3(n) &:= \sum_{i=n+1}^{\infty} C(i,1)(\pi^n(i) - \pi^{n-1}(i)). \end{aligned} \tag{3.7.18}$$

Recall now the assumption that the holding costs $C(n,1)$ and $C(n,0)$ are upper bounded by polynomials of finite degrees $P < \infty$ and $Q < \infty$, respectively. Hence, we can write $C(n,a) = E(n,a) + o(1)$, for large values of n, where $E(n,1) = \sum_{i=0}^{P} C^{(P,i)} n^i$, with $C^{(P,i)} := \lim_{n\to\infty} \frac{C(n,1) - \sum_{j=i+1}^{P} C^{(P,j)} n^j}{n^i}$, and $E(n,0) = \sum_{i=0}^{Q} E^{(Q,i)} n^i$, with $E^{(Q,i)} := \lim_{n\to\infty} \frac{C(n,0) - \sum_{j=i+1}^{Q} E^{(Q,j)} n^j}{n^i}$. We assume w.l.o.g. that P is such that $C^{(P,P)} > 0$ and Q such that $E^{(Q,Q)} > 0$. We then have

$$\begin{aligned} \xi_1(n) &= \sum_{i=1}^{n-1} E(i,0)(\pi^n(i) - \pi^{n-1}(i)) + o(1), \\ \xi_2(n) &= E(n,0)\pi^n(n) - E(n,1)\pi^{n-1}(n) + o(1), \\ \xi_3(n) &= \sum_{i=n+1}^{\infty} E(i,1)(\pi^n(i) - \pi^{n-1}(i)) + o(1). \end{aligned}$$

We now define $\hat{\xi}_1 := \sum_{i=1}^{n-1} E(i,0)(\pi^n(i) - \pi^{n-1}(i))$, and with the result obtained in Equation (3.7.15) we have that for large values of n $\hat{\xi}_1(n) \in \mathcal{O}\left(\frac{n^{Q-1}}{n!}\right) \subset o(1)$. Hence, for large values of n, $\xi_1(n) \in o(1)$. Let us now define $\hat{\xi}_2(n) := E(n,0)\pi^n(n) - E(n,1)\pi^{n-1}(n)$. Using (3.7.12) and (3.7.13) we have after some calculations,

$$\hat{\xi}_2(n) = \frac{\prod_{m=1}^{n} q^n(m-1,m)}{\prod_{m=1}^{n-1} q^n(m,m-1)} \left(\frac{E(n,0)\pi^n(0)}{q^n(n,n-1)} - \frac{E(n,1)\pi^{n-1}(0)}{q^{n-1}(n,n-1)} \right).$$

We recall that $q^{n-1}(n,n-1) = \mu + \theta' + \theta(n-1)$ and $q^n(n,n-1) = \theta n$, which together with (3.7.14) give, after some calculations,

$$\begin{aligned}\hat{\xi}_2(n) = &\prod_{m=1}^{n} \frac{q^n(m-1,m)}{q^n(m,m-1)} \frac{\theta n}{q^{n-1}(n,n-1)} \left(\left(E(n,0) - E(n,1)\right) \pi^n(0) + \mathcal{O}\left(\frac{n^{P-1}}{n!}\right)\right) \\ &+ \pi^n(n)(\mu + \theta' - \theta) \frac{E(n,0)}{q^{n-1}(n,n-1)}.\end{aligned}$$

Since for large values of n

$$\prod_{m=1}^{n} \frac{q^n(m-1,m)}{q^n(m,m-1)} \frac{\theta n}{q^{n-1}(n,n-1)} \cdot \mathcal{O}\left(\frac{n^{P-1}}{n!}\right) \subset \mathcal{O}\left(\frac{n^{P-1}}{(n!)^2}\right) \subset o(1),$$

we conclude that

$$\xi_2(n) = \frac{\pi^n(n)}{q^{n-1}(n,n-1)} \Big(\theta n (E(n,0) - E(n,1)) + (\mu + \theta' - \theta) E(n,0) \Big) + o(1). \tag{3.7.19}$$

Finally, we compute $\hat{\xi}_3(n) := \sum_{i=n+1}^{\infty} E(i,1)(\pi^n(i) - \pi^{n-1}(i))$. From (3.7.16) we see that

$$\hat{\xi}_3(n) = \frac{\mu + \theta' - \theta}{q^{n-1}(n,n-1)} \sum_{i=n+1}^{\infty} E(i,1)\pi^n(i) + \sum_{i=n+1}^{\infty} E(i,1) \cdot \mathcal{O}\left(\frac{1}{i!n!n}\right).$$

Since for large values of n $\sum_{i=n+1}^{\infty} E(i,1) \cdot \mathcal{O}\left(\frac{1}{i!n!n}\right) \subset \mathcal{O}\left(\frac{n^{P-1}}{i!n!}\right) \subset o(1)$, we obtain

$$\xi_3(n) = \frac{\mu + \theta' - \theta}{q^{n-1}(n,n-1)} \sum_{i=n+1}^{\infty} E(i,1)\pi^n(i) + o(1). \tag{3.7.20}$$

Now using $\xi_1 \in o(1)$, the expression of $\xi_2(n)$ in (3.7.19) and (3.7.20) and letting n be large, we see that $\frac{\xi_1(n)}{\pi^n(n)} \in o(1)$, and,

$$\begin{aligned}\frac{\xi_2(n)}{\pi^n(n)} &= \frac{\theta n (E(n,0) - E(n,1))}{\mu + \theta' + \theta(n-1)} + \frac{(\mu + \theta' - \theta) E(n,0)}{\mu + \theta' + \theta(n-1)} + o(1) \\ &= E(n,0) - E(n,1) + \frac{(\mu + \theta' - \theta)}{\theta n} E(n,0) + o(1) \\ &= E(n,0) - E(n,1) + \frac{(\mu + \theta' - \theta)}{\theta} \sum_{j=1}^{Q} E^{(P,j)} n^{j-1} + o(1),\end{aligned}$$

and

$$
\begin{aligned}
\frac{\xi_3(n)}{\pi^n(n)} =& \frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)} \cdot \sum_{i=n+1}^{\infty} E(i,1) \prod_{m=n+1}^{i} \frac{\lambda}{\mu+\theta'+\theta(m-1)} + o(1) \\
=& \frac{\mu+\theta'-\theta}{\theta n} \sum_{i=n+1}^{\infty} \sum_{j=0}^{P} C^{(P,j)} i^j \left(\frac{\lambda}{\theta m}\right)^{i-n} + o(1).
\end{aligned}
$$

Define $\tilde{W}^c(n)$ as $W^c(n)$ for large values of n. Substituting the expressions for $\xi_1(n)/\pi^n(n), \xi_2(n)/\pi^n(n)$ and $\xi_3(n)/\pi^n(n)$ in Equation (3.7.17), we obtain

$$
\begin{aligned}
&\tilde{W}^c(n) = (E(n,0) - E(n,1)) + (\mu+\theta'-\theta)/\theta \\
&\times \left(\sum_{j=1}^{Q} E^{(Q,j)} n^{j-1} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \right) + o(1),
\end{aligned}
$$

as $n \to \infty$, that is, the expression in Equation (3.3.4). $E(n,a)$ being non-decreasing together with Condition 3.1.2 implies that $\tilde{W}^c$ is non-decreasing, and hence W^∞ as well, which concludes the proof.

3.7.4 Proof of Propostion 3.5

For ease of notation we drop the dependency on k throughout the proof.

The index in the case $\mu+\theta' = \theta$ was obtained in (3.2.6), therefore we assume $\mu+\theta' > \theta$ throughout the proof. Recall from (3.3.3) that Whittle's index can be written as $\delta(\mu+\theta') - \delta'\theta' + W^c(n)$, where $W^c(n)$ corresponds to the holding costs only. Recall from (3.7.17) that $W^c(n)$ can be written as

$$
W^c(n) = \frac{\xi_1(n) + \xi_2(n) + \xi_3(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} \tag{3.7.21}
$$

with $\xi_i(n)$ for $i \in \{1,2,3\}$ as given in Equation (3.7.18).

Let us first compute $\lim_{\lambda\to 0} \pi^{n-1}(0)/\pi^n(0)$, since this result will be used later in the proof. Recall the expression of the steady-state probabilities as defined in (3.2.3). Using this together with (3.7.12) and (3.7.13) we obtain

$$
\begin{aligned}
\lim_{\lambda\to 0} \frac{\pi^{n-1}(0)}{\pi^n(0)} &= \lim_{\lambda\to 0} \frac{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} = 1 + \lim_{\lambda\to 0} \frac{\sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)} - \sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} \\
&= 1 + \lim_{\lambda\to 0} \frac{\sum_{m=n}^{\infty} \left(\frac{\lambda^m(\mu+\theta'+\theta(n-1))}{\theta n \prod_{i=1}^m q^{n-1}(i,i-1)} - \frac{\lambda^m \theta n}{\theta n \prod_{i=1}^m q^{n-1}(i,i-1)} \right)}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} = 1 + \frac{(\mu+\theta'-\theta)}{\theta n} \cdot \lim_{\lambda\to 0} \frac{\mathcal{O}(\lambda^n)}{1+\mathcal{O}(\lambda)} = 1.
\end{aligned}
\tag{3.7.22}
$$

Observe also the following

$$\lim_{\lambda\to 0}\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)} = \lim_{\lambda\to 0}\frac{\lambda^n/(\theta^n n!)}{-\frac{(\mu+\theta'-\theta)}{\theta n}\left(\frac{\frac{\lambda^n}{(\mu+\theta'+\theta(n-1))\theta^{n-1}(n-1)!}+\mathcal{O}(\lambda^{n+1})}{1+\mathcal{O}(\lambda)}\right)}$$
$$= \lim_{\lambda\to 0} -\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}+o(1) = -\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}. \tag{3.7.23}$$

Let us now consider the first term in (3.7.21), that is,

$$\frac{\sum_{m=0}^{n-1}C(m,0)(\pi^n(m)-\pi^{n-1}(m))}{\sum_{m=0}^{n}\pi^n(m)-\sum_{m=0}^{n-1}\pi^{n-1}(m)} = \frac{\sum_{m=0}^{n-1}C(m,0)\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}(\pi^n(0)-\pi^{n-1}(0))}{\pi^n(n)+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}(\pi^n(0)-\pi^{n-1}(0))}$$
$$=\frac{\sum_{m=0}^{n-1}C(m,0)\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}{\frac{\pi^n(n)}{\pi^n(0)-\pi^{n-1}(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}} = \frac{\sum_{m=0}^{n-1}C(m,0)\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}. \tag{3.7.24}$$

where in the first inequality we used the conditions (3.7.12) and (3.7.13). In order to obtain the limit of (3.7.24) as $\lambda\to 0$ we substitute the result obtained in (3.7.23), and we obtain the following

$$\lim_{\lambda\to 0}\frac{\xi_1(n)}{\sum_{m=0}^{n}\pi^n(m)-\sum_{m=0}^{n-1}\pi^{n-1}(m)} = \lim_{\lambda\to 0}\frac{\sum_{m=0}^{n-1}C(m,0)\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}$$
$$= \lim_{\lambda\to 0}\frac{C(0,0)+\mathcal{O}(\lambda)}{-\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}+1+\mathcal{O}(\lambda)} = -C(0,0)\frac{(\mu+\theta'-\theta)}{\theta n}. \tag{3.7.25}$$

Let us now consider the second term in (3.7.21), that is,

$$\frac{C(n,0)\pi^n(n)-C(n,1)\pi^{n-1}(n)}{\pi^n(n)+\sum_{m=0}^{n-1}\pi^n(n)-\sum_{m=0}^{n-1}\pi^n(n-1)} = \frac{C(n,0)-C(n,1)\frac{\pi^{n-1}(n)}{\pi^n(n)}}{1+\frac{1}{\pi^n(n)}(\pi^n(0)-\pi^{n-1}(0))\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}}$$
$$= \frac{C(n,0)-C(n,1)\frac{\theta n\pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)}}{1+\frac{\theta^n n!}{\lambda^n}(1-\pi^{n-1}(0)/\pi^n(0))\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}}. \tag{3.7.26}$$

Substituting the results obtained in (3.7.22) and (3.7.23) in the expression of Equation (3.7.26) we obtain

$$\lim_{\lambda\to 0}\frac{\xi_2(n)}{\sum_{m=0}^{n}\pi^n(n)-\sum_{m=0}^{n-1}\pi^{n-1}(m)} = \lim_{\lambda\to 0}\frac{C(n,0)-C(n,1)\left(\frac{\theta n}{\mu+\theta'+\theta(n-1)}\right)(1+\mathcal{O}(\lambda^n))}{1-\frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)}(1+\mathcal{O}(\lambda))}$$
$$= \lim_{\lambda\to 0}\frac{C(n,0)(\mu+\theta'+\theta(n-1))-C(n,1)\theta n+\mathcal{O}(\lambda^n)}{\theta n(1+\mathcal{O}(\lambda))}$$
$$= C(n,0)-C(n,1)+C(n,0)\frac{(\mu+\theta'-\theta)}{\theta n}+\mathcal{O}(\lambda). \tag{3.7.27}$$

To conclude the proof we need to analyze the third term in (3.7.21), that is,

$$\frac{\sum_{m=n+1}^{\infty} C(m,1)\pi^n(m) - \sum_{m=n+1}^{\infty} C(m,1)\pi^{n-1}(m)}{\pi^n(n) + \sum_{m=0}^{n-1} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)}$$
$$= \frac{\lambda^n \sum_{m=n+1}^{\infty} \frac{\lambda^{m-n}}{\prod_{i=1}^{n-1} q^n(i,i-1) \prod_{i=n+1}^{m} q^n(i,i-1)} \left(\frac{\pi^n(0)}{q^n(n,n-1)} - \frac{\pi^{n-1}(0)}{q^{n-1}(n,n-1)} \right)}{\lambda^n \left(\frac{\pi^n(0)}{\theta^n n!} + \frac{1}{\lambda^n} \left(\pi^n(0) - \pi^{n-1}(0)\right) \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m} \right)}$$
$$= \frac{\sum_{m=n+1}^{\infty} \frac{\lambda^{m-n}}{\prod_{i=n+1}^{m} q^n(i,i-1)} \left(1 - \frac{\theta n \pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)} \right)}{\left(1 + \frac{\theta^n n!}{\lambda^n} \left(1 - \frac{\pi^{n-1}(0)}{\pi^n(0)} \right) \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m} \right)}.$$

In the last expression we substitute the results obtained in (3.7.22) and (3.7.23), and we show that

$$\lim_{\lambda\to 0} \frac{\xi_3(n)}{\sum_{m=0}^{n} \pi^n(n) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} = \lim_{\lambda\to 0} \frac{\sum_{m=n+1}^{\infty} \frac{\lambda^{m-n}}{\prod_{i=n+1}^{m} q^n(i,i-1)} \left(1 - \frac{\theta n \pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)} \right)}{\left(1 + \frac{\theta^n n!}{\lambda^n} \left(1 - \frac{\pi^{n-1}(0)}{\pi^n(0)} \right) \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m} \right)}$$
$$= \lim_{\lambda\to 0} \frac{\sum_{m=n+1}^{\infty} \frac{\lambda^{m-n}}{\prod_{i=n+1}^{m} q^n(i,i-1)} \left(1 - \frac{\theta n}{\mu+\theta'+\theta(n-1)} \left(1 + \mathcal{O}(\lambda^n)\right) \right)}{\left(1 - \left(\frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)} + \mathcal{O}(\lambda) \right) \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m} \right)} = \lim_{\lambda\to 0} \frac{\mathcal{O}(\lambda)}{\frac{\theta n}{\mu+\theta'+\theta(n-1)} + \mathcal{O}(\lambda)} = 0. \tag{3.7.28}$$

We now substitute the results obtained in Equations (3.7.25), (3.7.27) and (3.7.28) in $\lim_{\lambda\to 0} W^c(n)$, and we obtain

$$\lim_{\lambda\to 0} W^c(n) = C(n,0) - C(n,1) + \frac{(\mu+\theta'-\theta)}{\theta n}(C(n,0) - C(0,0)).$$

3.7.5 Proof of Proposition 3.6

For ease of notation we drop the dependency on k throughout the proof.

The index in the case $\mu+\theta' = \theta$ was obtained in (3.2.6), therefore we assume $\mu+\theta' > \theta$ throughout the proof. Recall from (3.3.3) that Whittle's index can be written as $\delta(\mu+\theta') - \delta'\theta' + W^c(n)$, where $W^c(n)$ corresponds to the holding costs only. Recall from (3.7.17) that $W^c(n)$ can be written as

$$W^c(n) = \frac{\xi_1(n) + \xi_2(n) + \xi_3(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} \tag{3.7.29}$$

with $\xi_i(n)$ for $i \in 1,2,3$ as given by Equation (3.7.18)

We first compute $\pi^{n-1}(0)/\pi^n(0)$, which will be used later in the proof;

$$\frac{\pi^{n-1}(0)}{\pi^n(0)} = \frac{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} = \left(1 + \frac{\sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)} - \sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}\right)$$

$$= 1 + \frac{(\mu + \theta' - \theta)}{\theta n} \cdot \frac{\sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} = 1 + \frac{(\mu + \theta' - \theta)}{\theta n} \cdot \frac{1}{1 + \frac{\sum_{m=0}^{n-1} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}} \tag{3.7.30}$$

$$= 1 + \frac{\mu + \theta' - \theta}{\theta n}(1 + o(1)). \tag{3.7.31}$$

The latter inequality is obtained by letting $\lambda \to \infty$. We now proceed to compute (3.7.29) as $\lambda \to \infty$. Let us begin by computing the term that corresponds to $\xi_1(n)$. We have after some algebra

$$\begin{aligned}\frac{\xi_1(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} &= \frac{\sum_{m=0}^{n-1} C(m,0)(\pi^n(m) - \pi^{n-1}(m))}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} \\ &= \frac{\sum_{m=0}^{n-1} C(m,0)\frac{\lambda^m}{\theta^m m!}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)} + \sum_{m=0}^{n-1} \frac{\lambda^m}{\theta^m m!}},\end{aligned} \tag{3.7.32}$$

which after substitution of (3.7.31) reduces to

$$\frac{\xi_1(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = \mathcal{O}\left(\frac{1}{\lambda}\right), \tag{3.7.33}$$

as $\lambda \uparrow \infty$, for all n. We are now interested in computing the second term in (3.7.29) as $\lambda \to \infty$. Using (3.7.31) we obtain

$$\begin{aligned}\lim_{\lambda\to\infty} \frac{\xi_2(n)}{\sum_{m=0}^{n} \pi^n(n) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} &= \lim_{\lambda\to\infty} \frac{C(n,0) - C(n,1)\frac{\pi^{n-1}(n)}{\pi^n(n)}}{1 + \frac{\pi^n(0)-\pi^{n-1}(0)}{\pi^n(n)} \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}} \\ = \lim_{\lambda\to\infty} \frac{C(n,0) - C(n,1)\frac{\theta n}{\mu+\theta'+\theta(n-1)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}{1 + \frac{1-\pi^{n-1}(0)/\pi^n(0)}{\lambda^n/(\theta^n n!)} \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}} &= C(n,0) - C(n,1),\end{aligned} \tag{3.7.34}$$

for all n. We are left with the third term in (3.7.29), that is,

$$\begin{aligned}&\frac{\xi_3(n)}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} \\ &= \frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{n-1} q^n(i,i-1)\prod_{i=n+1}^{m} q^n(i,i-1)}\left(\frac{\pi^n(0)}{\theta n} - \frac{\pi^{n-1}(0)}{\mu+\theta'+\theta(n-1)}\right)}{\pi^n(n) + (\pi^n(0) - \pi^{n-1}(0))\sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}} \\ &= \frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{m} q^n(i,i-1)}\left(\frac{\theta n}{\mu+\theta'+\theta(n-1)}\left(1 - \frac{\pi^{n-1}(0)}{\pi^n(0)}\right) + \frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)}\right)}{\lambda^n/(\theta^n n!) + (1 - \pi^{n-1}(0)/\pi^n(0))\sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}},\end{aligned} \tag{3.7.35}$$

where in the second step we used that $\prod_{i=1}^{n-1} q^n(i,i-1)\prod_{i=n+1}^{m} q^n(i,i-1) = \prod_{i=1}^{m} q^n(i,i-1)/\theta n$. After substituting (3.7.30) in the latter equation and some algebra, we obtain that (3.7.35) can be written as

$$\frac{\mu+\theta'-\theta}{\theta n}\frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}}{\frac{\lambda}{\theta n}\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}(1+o(1))}.$$

Hence the third term as $\lambda\to\infty$ simplifies to

$$\frac{\mu+\theta'-\theta}{\theta}\frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}}{\frac{\lambda}{\theta}\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}} + o(1) = \frac{\mu+\theta'-\theta}{\theta}\frac{\sum_{m=n+1}^{\infty} C(m,1)\pi^{n-1}(m)}{\lambda/\theta} + o(1). \tag{3.7.36}$$

The latter equality follows due to $\pi^{n-1}(0) = (\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)})^{-1}$. We now write (3.7.36) as follows

$$\begin{aligned}
&\frac{\mu+\theta'-\theta}{\theta}\left(\frac{\sum_{m=0}^{\infty} C(m,1)\pi^{n-1}(m)}{\lambda/\theta} - \frac{\sum_{m=0}^{n} C(m,1)\pi^{n-1}(m)}{\lambda/\theta}\right) + o(1)\\
&= \frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta}\left(1 - \frac{\sum_{m=0}^{n} C(m,1)\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}{\mathcal{O}(\lambda^n) + \sum_{m=n+1}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}\right) + o(1),
\end{aligned} \tag{3.7.37}$$

where

$$\mathbb{E}(C(N^{n-1},1)) = \frac{\sum_{m=0}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}.$$

We then have that if there exists $z\geq 1$ such that $\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda^z}\to 0$, as $\lambda\to\infty$, then (3.7.37) reduces to

$$\frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta} + o(1),$$

Hence, together with Equations (3.7.29), (3.7.33) and (3.7.34) we obtain that

$$W^c(n) = C(n,0) - C(n,1) + \frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta} + o(1),$$

as $\lambda\to\infty$. This concludes the proof.

3.7.6 Proof of Proposition 3.7

For ease of notation, we omit the class index k in the proof.

Since $\theta' = \theta$ we have $\mu + \theta' > \theta$. Since $\delta' = \delta = 0$, $\theta' = \theta$ and $C(n, a) = C(n)$, we can write $\tilde{C}(n, a) = C(n)$. Hence, we are interested in the following limit

$$\begin{aligned}\lim_{\theta\to 0} \theta W(n) &= \lim_{\theta\to 0} \frac{\theta \sum_{m=0}^{\infty} C(m)\left(\pi^n(m) - \pi^{n-1}(m)\right)}{\sum_{m=1}^{n-1}\left(\pi^n(m) - \pi^{n-1}(m)\right) + \pi^n(n)} \\ &= \varepsilon_1(n)\varepsilon_2(n),\end{aligned}$$

with

$$\varepsilon_1(n) = \lim_{\theta\to 0} \frac{\theta}{\sum_{m=1}^{n-1}\left(\pi^n(m) - \pi^{n-1}(m)\right) + \pi^n(n)},$$

and

$$\varepsilon_2(n) = \lim_{\theta\to 0} \sum_{m=0}^{\infty} C(m)\left(\pi^n(m) - \pi^{n-1}(m)\right).$$

Consider $\varepsilon_2(n)$. We have

$$\pi^n(0) - \pi^{n-1}(0) \xrightarrow{\theta\to 0} 0.$$

hence

$$\pi^n(m) - \pi^{n-1}(m) \xrightarrow{\theta\to 0} 0, \qquad \forall m < n-1,$$

$$\pi^n(n-1) - \pi^{n-1}(n-1) \xrightarrow{\theta\to 0} (\rho - 1),$$

and

$$\pi^n(m) - \pi^{n-1}(m) \xrightarrow{\theta\to 0} \rho^{m-n}(1-\rho)^2, \qquad \forall m \geq n.$$

It can be proven that $|\pi^n(m) - \pi^{n-1}(m)| < \rho^{m-n}\kappa$ for all n and $m \geq n$, with κ a θ-independent constant (this is done at the end of this proof). By assumption $C(\cdot)$ is bounded by a polynomial of finite degree. Together this implies that the infinite sum in $\varepsilon_2(n)$ is dominated by $\sum_{m=0}^{\infty} C(m)\rho^{m-n}$ which converges for $\rho < 1$. Therefore, due to the dominated convergence theorem, the limit and the summation can be interchanged in $\varepsilon_2(n)$, and one obtains

$$\begin{aligned}\varepsilon_2(n) &= -C(n-1)(1-\rho) + \frac{(1-\rho)}{\rho}\sum_{m=n}^{\infty} C(m)(1-\rho)\rho^{m-n+1} \\ &= \frac{(1-\rho)}{\rho}\left(-C(n-1) + \sum_{m=0}^{\infty} C(m+n-1)(1-\rho)\rho^m\right).\end{aligned}$$

After some algebra and using that $\pi^n(n) \xrightarrow{\theta\to 0} (1-\rho)$ (as pointed out in Section 3.4), we obtain $\varepsilon_1(n) = 1/\mu$.

To conclude the proof we need to prove that $|\pi^n(m) - \pi^{n-1}(m)| < \rho^{m-n}\kappa$ for all n and $m \geq n$, with κ constant. By the definition of $\pi^n(m)$ we have

$$\begin{aligned}|\pi^n(m) - \pi^{n-1}(m)| &= \left|\frac{\lambda^m}{\theta^{n-1}(n-1)!\prod_{j=n+1}^{m}(\mu+\theta j)}\left(\frac{\pi^n(0)}{\theta n} - \frac{\pi^{n-1}(0)}{\mu+\theta n}\right)\right|, \\ &\leq \rho^{m-n}\frac{\lambda^n}{\theta^{n-1}(n-1)!}\left|\left(\frac{\pi^n(0)}{\theta n} - \frac{\pi^{n-1}(0)}{\mu+\theta n}\right)\right|,\end{aligned}$$

where the inequality holds due to $\frac{\lambda^{m-n}}{\prod_{j=n+1}^{m}(\mu+\theta j)} \leq \left(\frac{\lambda}{\mu}\right)^{m-n} = \rho^{m-n}$, for all $\theta \geq 0$. It then suffices to prove that $\frac{\lambda^n}{\theta^{n-1}(n-1)!}\left|\left(\frac{\pi^n(0)}{\theta n} - \frac{\pi^{n-1}(0)}{\mu+\theta n}\right)\right|$ is bounded by some constant κ. Substituting the values of $\pi^n(0)$ and $\pi^{n-1}(0)$, we see that

$$\begin{aligned}\left|\frac{\pi^n(0)}{\theta n} - \frac{\pi^{n-1}(0)}{\mu+\theta n}\right| &= \left|\frac{1}{\theta n(\sum_{m=0}^{n}\frac{\lambda^m}{\theta^m m!} + \sum_{m=n+1}^{\infty}\frac{\lambda^m}{\theta^n n!\prod_{j=n+1}^{m}(\mu+\theta j)})}\right. \\ &\quad \left. - \frac{1}{(\mu+\theta n)(\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!} + \sum_{m=n}^{\infty}\frac{\lambda^m}{\theta^{n-1}(n-1)!\prod_{j=n+1}^{m}(\mu+\theta j)})}\right| \\ &= \frac{\mu\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}}{\theta n(\mu+\theta n)}\pi^n(0)\pi^{n-1}(0) \leq \frac{1}{\theta n}\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}\pi^n(0)\pi^{n-1}(0).\end{aligned}$$

It then suffices to show that

$$\frac{\lambda^n}{\theta^n n!}\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}\pi^n(0)\pi^{n-1}(0) = \pi^n(n)\cdot\sum_{m=0}^{n-1}\pi^{n-1}(m),$$

is bounded by κ, a θ-independent constant. To obtain the equality we used that $\pi^n(m) = \lambda^m\pi^n(0)/\theta^m m!$ for all $m \leq n$. The RHS of the equality being bounded by a constant is obvious since $\pi^n(n) < 1$ and $\sum_{m=0}^{n-1}\pi^{n-1}(m) < 1$ for all $\theta \geq 0$, therefore $\kappa = 1$.

This concludes the proof.

3.7.7 Proof of Proposition 3.8

We first assume there exists a k such that $C_k(0,1) > 0$. Let us consider that $W = 0$, and from (2.5.1) we know that necessarily $\mathcal{C}^{REL(0)}(0) \leq \mathcal{C}^{OPT}$. We also consider the policy $\bar{u} \in \mathcal{U}$ that takes active action when the total number of customers in the system is 0, and is passive otherwise. Note that policy $\bar{u}$ does not take any *scheduling* decision. Since $\mu_k + \theta'_k \geq \theta_k$, for all k, the queue length under policy $\bar{u}$ stochastically upper bounds any policy $u \in \mathcal{U}$. Note that under the assumption $C_k(0,0) \geq C_k(0,1)$, $\forall k$, it holds from (3.1.2) that, for all n, $C_k(n,0) \geq C_k(n,1)$, which implies that $W_k(n)$ is always positive, see Section 3.2.3. Hence, it follows $\mathcal{C}^{WI} \leq \mathcal{C}^{\bar{u}}$. We will now show that $\frac{\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}}{\mathcal{C}^{OPT}} \to 0$ as $\lambda \to 0$, which in view of (2.5.1) implies the optimality of Whittle's index policy.

We have $W_k(0) = C_k(0,0) - C_k(0,1) \geq 0$, for all k. Setting $W = 0$, it follows that for every class $REL(0)$ is the threshold policy with threshold -1, that is, class-k always activates for any state $n_k > -1$. Hence, under policy $REL(0)$ the steady-state probabilities for class-k are given by (3.2.3) with threshold $n = -1$. It then follows that

$$\begin{aligned}\mathcal{C}^{REL(0)}(0) &= \sum_{k=1}^{K}\sum_{m=0}^{\infty} C_k(m,1)\pi_k^{-1}(m) \\ &= \sum_{k=1}^{K} C_k(0,1)\pi_k^{-1}(0) + \sum_{k=1}^{K} C_k(1,1)\frac{\lambda\gamma_k}{\mu_k+\theta'_k}\pi_k^{-1}(0) + \mathcal{O}(\lambda^2),\end{aligned} \tag{3.7.38}$$

as $\lambda \downarrow 0$. We have $\pi_k^{-1}(0) = (1+\mathcal{O}(\lambda))^{-1}$, hence $\mathcal{C}^{REL(0)}(0) = \sum_{k=1}^{K} C_k(0,1) + \mathcal{O}(\lambda)$.

Under policy $\bar{u} \in \mathcal{U}$, every class k behaves as an $M/M/\infty$ queue with arrival rate $\lambda\gamma_k$ and departure rate $\theta_k n_k$. We then have $\mathcal{C}^{\bar{u}} = \sum_{k=1}^{K} C_k(0,1)\mathrm{e}^{-\lambda\gamma_k/\theta_k} + \sum_{k=1}^{K}\sum_{m=1}^{\infty} C_k(m,0)\frac{(\lambda\gamma_k)^m}{\theta_k^m m!}\mathrm{e}^{-\lambda\gamma_k/\theta_k} = \sum_{k=1}^{K} C_k(0,1) + \mathcal{O}(\lambda)$.

Hence,

$$\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}(0) \quad = \quad \mathcal{O}(\lambda). \tag{3.7.39}$$

We now note that in the limit $\lambda \to 0$, $\mathcal{C}^{OPT} \geq \mathcal{C}^{REL(0)}(0) = \mathcal{O}(1)$. Together with (2.5.1) and (3.7.39), we thus conclude that

$$\lim_{\lambda\downarrow 0} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} \leq \lim_{\lambda\downarrow 0} \frac{\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}(0)}{\mathcal{C}^{OPT}} = 0.$$

In case $C_k(0,1) = 0$, then $C_k(0,0) \geq C_k(0,1)$ for all k and $W_k(0) = C_k(0,0) - C_k(0,1) \geq 0$, for all k. Setting $W = 0$, it follows that $REL(0)$ is the policy that activates class-k for any $n_k \geq 0$. We consider $\bar{u}$ to be the policy that takes active action when the total number of customers in the system is 0 or 1, and is passive otherwise. Then

$$\mathcal{C}^{\bar{u}} = \sum_{k=1}^{K} C_k(1,1)\frac{\lambda\gamma_k}{\mu_k + \theta'_k}\pi_k^{\bar{u}}(0) + \sum_{k=1}^{K}\sum_{m=2}^{\infty} C_k(m,0)\frac{(\lambda\gamma_k)^m}{(\mu_k + \theta'_k)\theta^{m-1}m!}\pi_k^{\bar{u}}(0),$$

and $\pi_k^{\bar{u}}(0) = \left(1 + \frac{\lambda\gamma_k}{\mu_k+\theta'_k} + \frac{(\lambda\gamma_k)^2}{(\mu_k+\theta'_k)2\theta_k} + \mathcal{O}(\lambda^3)\right)^{-1}$ as $\lambda \to 0$. We have that $\pi_k^{-1}(0) = \pi_k^{\bar{u}}(0) + \mathcal{O}(\lambda^2)$ as $\lambda \to 0$. Then the term that corresponds to $C(1,1)$ in $\mathcal{C}^{\bar{u}}$ and $\mathcal{C}^{REL(0)}(0)$ as given in (3.7.38) coincide up to a $\mathcal{O}(\lambda^2)$ term. Hence, $\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}(0) = \mathcal{O}(\lambda^2)$ and $\mathcal{C}^{OPT} \geq \mathcal{C}^{REL(0)}(0) = \mathcal{O}(\lambda)$, which lead to the desired result $\lim_{\lambda\downarrow 0} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} = 0$.

3.7.8 Proof of Proposition 3.9

We have assumed that Whittle's index is constant, that is, $W_k(n) = w_k$, and that there exists $\bar{k} \in \{1, \ldots, K\}$ such that $w_{\bar{k}} > w_k(n)$ for all n and all $k \neq \bar{k}$.

We prove Proposition 3.9 as follows:

- **Step 1:** We prove that there exists $\bar{W}$ such that

$$\lim_{\lambda\to\infty} \frac{\mathcal{C}^{REL(\bar{W})}(\bar{W})}{\mathcal{C}^{REL(\bar{W})}} = 1.$$

- **Step 2:** We note that for the choice of $\bar{W}$ in Step 1, $\mathcal{C}^{WI} \leq \mathcal{C}^{REL(\bar{W})}$. Then applying the result in Equation (2.5.1) we obtain $\lim_{\lambda\to\infty} \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} = 1$.

- **Step 3:** Steps 1 and 2 and the result in Equation (2.5.1) imply $\lim_{\lambda\to\infty} \frac{\mathcal{C}^{OPT}}{\mathcal{C}^{WI}} = 1$.

Step 1. From the assumption in the statement $W_k(n_k) = w_k$ with w_k constant for all k, and $w_{\bar{k}} > w_k$ for all $k \neq \bar{k}$. We can therefore find $\bar{W}$ constant such that $w_k < \bar{W} < w_{\bar{k}}$ for all $k \neq \bar{k}$. We will then prove that $\lim_{\lambda\uparrow\infty} \mathcal{C}^{REL(\bar{W})}(\bar{W})/\mathcal{C}^{REL(\bar{W})} = 1$. Recall from (2.3.2) that

$$\mathcal{C}^{REL(\bar{W})}(\bar{W}) = \sum_{k=1}^{K} \mathbb{E}(\tilde{C}(N_k^{REL(\bar{W})}, S_k^{REL(\bar{W})}(N_k^{REL(\bar{W})}))) \\ - \bar{W}\left(1 - K + \sum_{k=1}^{K}\left(1 - \mathbb{E}\left(S_k^{REL(\bar{W})}(N_k^{REL(\bar{W})})\right)\right)\right).$$

Since by assumption $w_k < w_{\bar{k}}$ for all $k \neq \bar{k}$, we have that the policy $REL(\bar{W})$ will never serve class k (with $k \neq \bar{k}$) and will serve class $\bar{k}$ as long as there is a class-$\bar{k}$ customer in the system. One can therefore prove

$$\sum_{k=1}^{K} \mathbb{E}(S^{REL(\bar{W})}(N_k^{REL(\bar{W})}) = 1) \to 1,$$

as $\lambda \uparrow \infty$, as in heavy traffic there are always class-$\bar{k}$ customers. Since

$$\mathcal{C}^{REL(\bar{W})} = \sum_{k=1}^{K} \mathbb{E}(\tilde{C}(N_k^{REL(\bar{W})}, S^{REL(\bar{W})}(N_k^{REL(\bar{W})}))),$$

the latter implies

$$\lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}(\bar{W})}{\mathcal{C}^{REL(\bar{W})}} = 1.$$

Step 2. We now prove that $\lim_{\lambda\uparrow\infty} \mathcal{C}^{REL(\bar{W})}/\mathcal{C}^{WI} = 1$. Note that Whittle's index policy and the relaxed policy behave the same whenever $N_{\bar{k}} > 0$. When $N_{\bar{k}} = 0$ WI serves class-k customers (with $k \neq \bar{k}$) while $REL(\bar{W})$ idles. Since $\mu_k + \theta'_k \geq \theta_k$ and $C_k(m, 0) \geq C_k(m, 1)$ for all k, it follows directly that $N_k^{WI} \leq_{st} N_k^{REL}$ for all $k \neq \bar{k}$ while $N_{\bar{k}}^{WI}$ and $N_{\bar{k}}^{REL(\bar{W})}$ are stochastically identical. This implies $\mathcal{C}^{WI} \leq \mathcal{C}^{REL(\bar{W})}$, and therefore $1 \leq \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}}$. From (2.5.1), we have

$$\mathcal{C}^{REL(\bar{W})}(\bar{W}) \leq \mathcal{C}^{WI} \Longrightarrow \frac{\mathcal{C}^{REL(\bar{W})}(\bar{W})}{\mathcal{C}^{REL(\bar{W})}} \cdot \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} \leq \frac{\mathcal{C}^{WI}}{\mathcal{C}^{WI}} = 1 \Longrightarrow \lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} \leq 1.$$

In the last implication the result in Step 1 has been used. We therefore have $\lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} \leq 1$ and $1 \leq \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}}$ which prove that $\lim_{\lambda\to\infty} \mathcal{C}^{REL(\bar{W})}/\mathcal{C}^{WI} = 1$.

Step 3. From Equation (2.5.1) we obtain

$$\lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}(\bar{W})}{\mathcal{C}^{REL(\bar{W})}} \cdot \lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} \leq \lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{OPT}}{\mathcal{C}^{WI}} \leq \frac{\mathcal{C}^{WI}}{\mathcal{C}^{WI}} = 1.$$

The results in Steps 1 and 2 then imply $\mathcal{C}^{OPT}/\mathcal{C}^{WI} \to 1$ as $\lambda \to \infty$, which concludes the proof.

Part II

Dynamic control of fluid resource-sharing systems

Chapter 4

Fluid index policy for birth-and-death restless bandits

Contents

4.1 Fluid version of relaxed optimization problem **72**

4.2 Equivalence of stochastic index and fluid index in light-traffic regime **76**

4.3 Case studies **78**

4.4 Appendix **90**

In Part I we have seen how to obtain efficient index policies for RBP with birth-and-death bandit's state evolution. Whittle's index can be efficiently calculated, as we have done it for a multi-class abandonment queue in Chapter 3. However, in general it requires several technical conditions to be verified and only for particular instances it provides qualitative insights on how Whittle's index policy depends on the input parameters. We therefore formulate a fluid version of the relaxed optimization problem. This approach is motivated by the pioneering work of Avram *et al.* [14] and Weiss [97] where fluid control models were used to approximate stochastic optimization problems. One of the advantages of this approximation is that one can solve the unconstrained deterministic optimization problem to obtain an explicit characterization of an optimal control. This approach provides an explicit index, which we will denote by the fluid index. We then develop the analogous theory of Whittle's index and propose the fluid index policy as a heuristic. The latter *does* provide qualitative insights and is close to Whittle's index. We apply Whittle's index policy and the fluid index policy to several problems: *e.g.*, a multi-class single-server abandonment queue, opportunistic scheduling in wireless systems, power-aware server-farms, and inventory management with perishable item. Numerical simulations show that our fluid index policy is nearly optimal.

The chapter is organized as follows. In Section 4.1 we present the fluid approximation of the relaxed optimization problem introduced in Chapter 2. The latter allows us to derive a fluid index, which enables to define a fluid index policy. In Section 4.2 we establish equivalences between the fluid index and the Whittle index. Finally, in Section 4.3 we apply the fluid index policy to the four problems mentioned above. Most proofs can be found in Appendix 4.4.

4.1 Fluid version of relaxed optimization problem

In this section we will solve the fluid version of the unidimensional unconstrained optimization problem (2.3.3), that is, we only take into account the average behavior of the original stochastic system. In Section 4.1.1 we describe the fluid dynamics and the fluid version of problem (2.3.3). In Section 4.1.2 we give the solution of the unconstrained fluid model and the fluid index. In Section 4.1.3 we define the fluid index policy, which serves as a heuristic for the original problem.

4.1.1 Fluid model and bias optimality

We approximate the stochastic unconstrained optimization problem as presented in Section 2.3 by a deterministic fluid model, where bandit k has a continuous state space $[0,\infty)$ instead of a discrete state space $\{0,1,\ldots\}$. The fluid dynamics will be defined by only taking into account the mean dynamics of the stochastic process.

Let $m_k(t)\in[0,\infty)$ be the state of bandit k and $s_k(t)\in\{0,1\}$ the control parameter. Let u denote a fluid control that determines $s_k^u(t)$, that is, whether bandit k is active or not. We use the following compact notation for the drift under action a:

$$f_k^a(m_k) := b_k^a(m_k) - d_k^a(m_k), \quad a=0,1,$$

with $m_k\geq 0$, where for non-integer values of m_k the functions b_k^0, d_k^0, b_k^1 and d_k^1 are defined such that they are continuous. We further assume $f_k^a(m_k)$ to be non-increasing in m_k for $a\in\{0,1\}$. The fluid dynamics under control u can then be written as follows:

$$\frac{\mathrm{d}m_k^u(t)}{\mathrm{d}t} =(1-s_k^u(t))f_k^0(m_k^u(t)) + s_k^u(t)f_k^1(m_k^u(t)), \tag{4.1.1}$$

where the control u is such that $m_k^u(t)\geq 0$ for all t.

At time t, we define the cost for the fluid version of the unconstrained problem (2.3.3) as

$$C_k(m_k(t),s_k(t)) := (1-s_k(t))C_k(m_k(t),0) + s_k(t)C(m_k(t),1),$$

where in non-integer values for m_k the function $C_k(m_k,a)$ is defined such that it is continuous in m_k. We assume the function $C_k(m_k,a)$ to be convex in m_k for $a=0,1$.

An *equilibrium point* $(\bar{m}_k,\bar{s}_k)$ of the fluid dynamics is such that $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t}=0$, that is, $(1-\bar{s}_k)f_k^0(\bar{m}_k)+\bar{s}_k f_k^1(\bar{m}_k)=0$, with $\bar{s}_k\in[0,1]$. That is, in equilibrium, a fraction of time $\bar{s}_k$ $(1-\bar{s}_k)$ the action $a=1$ $(a=0)$ is chosen. Define $\bar{s}_k(\bar{m}_k):=f_k^0(\bar{m}_k)/(f_k^0(\bar{m}_k)-f_k^1(\bar{m}_k))$ and we assume throughout this chapter $\bar{s}_k(\bar{m}_k)$ to be strictly monotone in $\bar{m}_k$. A discussion on the latter assumption can be found in Remark 4.2.

In the stochastic model we aim to minimize for a given bandit the relaxed optimization problem, that is, we minimize the time-average of the cost minus the subsidy obtained, as stated in (2.3.3). In equilibrium, $\bar{s}_k$ is the average amount of time the system is active, hence, the fluid version of (2.3.3) will be to find the equilibrium point that minimizes the cost at equilibrium $EC_k(\bar{s},W)$, where

$$EC_k(\bar{s}_k,W) := (1-\bar{s}_k)C_k(\bar{m}_k,0) + \bar{s}_k C(\bar{m}_k,1) - W(1-\bar{s}_k).$$

We denote by (m_k^*, s_k^*) an optimal equilibrium point and define the optimal equilibrium cost under subsidy W as

$$EC_k^*(W) := (1 - s_k^*)(C_k(m_k^*, 0) - W) + s_k^* C_k(m_k^*, 1). \tag{4.1.2}$$

Since the time-average criteria will be attained by any control that has m_k^* as equilibrium point, we are interested in controls that are *bias-optimal.* That is, among all controls that reach the optimal equilibrium point, a bias-optimal control is the one that minimizes the cost to get to this equilibrium point. Hence, our aim is to find the control u that minimizes the total bias cost, that is, the cost and subsidy obtained over time minus the optimal cost in equilibrium, denoted as

$$J_k^u(m_k(0), W) := \int_0^\infty \big(C_k(m_k(t), s_k^u(t)) - W(1 - s_k^u(t)) - EC_k^*(W)\big)\mathrm{d}t. \tag{4.1.3}$$

We further define $J_k(m_k(0), W) := \min_u J_k^u(m_k(0), W)$.

The theory of optimal control shows that a sufficient condition in order for a control to be bias-optimal is to solve the HJB equation, see Section 1.3.4:

$$EC_k^*(W) = \min\Big(C_k(m_k, 1) + f_k^1(m_k)\frac{\partial J_k(m_k, W)}{\partial m_k}, C_k(m_k, 0) - W + f_k^0(m_k)\frac{\partial J_k(m_k, W)}{\partial m_k}\Big). \tag{4.1.4}$$

Then, for a given state m_k, an optimal action in that state is given by a minimizer of the right-hand-side in (4.1.4).

The main advantage of our approach is that (4.1.4) can be solved in general, see Proposition 4.1, while solving (2.3.3) (or equivalently (2.3.4)) requires to establish that an optimal policy for the relaxed problem is of threshold structure.

Remark 4.1. *An alternative route to obtain* (4.1.3) *is to consider the total discounted cost criterion*

$$\mathcal{C}^\phi(\beta) := \sum_{k=1}^K \mathbb{E}\left(\int_0^\infty \mathrm{e}^{-\beta t} C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t)))\ dt\right),$$

with $\beta > 0$ a discount factor, and to consider its fluid version. We then get a deterministic control problem under the total discounted cost criterion which is difficult to solve in general. As in Section 2.3, we relax the service constraint and allow that the total discounted number of bandits active is $M/(1-\beta)$ or lower. For a single bandit, the objective of the unconstrained fluid problem with discounted cost is then to find a control u that minimizes $J_k^{u,\beta}(m_k(0), W) := \int_0^\infty \mathrm{e}^{-\beta t}\big(\big(C_k(m_k(t), s_k^u(t)) - W(1 - s_k^u(t))\big)\big)dt$. Hence, an optimal control for the unconstrained fluid discounted control problem is the solution of

$$\begin{aligned}\beta J_k^\beta(m_k, W) = \min_s(&C_k(m_k, 1) + \beta f_k^1(m_k)\partial J_k^\beta(m_k, W)/\partial m_k,\\ &C_k(m_k, 0) - W + \beta f_k^0(m_k)\partial J_k^\beta(m_k, W)/\partial m_k),\end{aligned} \tag{4.1.5}$$

see Puterman [77, Chapter 10], where $J_k^\beta(m_k, W) = \min_u J_k^{u,\beta}(m_k, W)$. We now note that as $\beta \to 0$, $\beta J_k^\beta(m_k, W) \to EC_k^*(W)$, see Puterman [77, Corollary 8.2.5], and we thus obtain that (4.1.5) converges to (4.1.4).

Remark 4.2. *As highlighted in Section 3.4 the assumption on $\sum_{m=0}^n \pi_k^n(m)$ being strictly increasing (required for indexability of Whittle's index) excludes certain models. In the fluid context some of the* *models are excluded from our analysis due to the assumption that $\bar{s}_k(\bar{m}_k)$ is strictly monotone in $\bar{m}_k$. Let us consider an M/M/1 queue with controlled arrivals, where arrivals are exponentially distributed with rate $\mu_k a$, $a = 0, 1$ and departures follow a Poisson process with rate $\mu_k < \lambda_k$. Then $(1-\bar{s}_k)f_k^0(m_k)+\bar{s}_k f_k^1(m_k) = -\mu_k + \bar{s}_k\lambda_k$. The latter equals 0 when $\bar{s}_k = \mu_k/\lambda_k$. Hence, when $\bar{s}_k = \mu_k/\lambda_k$, any m_k is an equilibrium point. The latter implies that the fraction of time the system is passive is the same no matter the equilibrium considered. Hence, the subsidy for passivity does not discriminate between states. In Section 4.3 we overcome this issue by considering items to be perishable in the make-to-stock problem.*

4.1.2 Optimal fluid control and fluid index

In this section we derive an optimal solution for the unconstrained fluid problem (4.1.3) for a given bandit. This solution is described by a fluid index function, which allows a simple closed-form expression. Based on the fluid index we define in Section 4.1.3 a heuristic for the original stochastic model, which we will show in Section 4.3 to perform nearly optimal.

In order to define the fluid index, we need the following notation: we denote by m_k^a the value of m_k such that $f_k^a(m_k) = 0$, $a = 0, 1$. We adopt the convention that $m_k^a = \infty$ in case $f_k^a(m_k) > 0$ for all $m_k \geq 0$, and that $m_k^a = 0$ in case $f_k^a(m_k) < 0$ for all $m_k \geq 0$, that is, $m_k^a \in [0, \infty)$. The structure of the fluid index will depend on how m_k^1 and m_k^0 are ordered. In Figure 4.1 we show the drifts in case $m_k^1 < m_k^0$. The shape of the fluid index depends on whether the state m_k is such that $m_k < m_k^1$, $m_k \in [m_k^1, m_k^0]$, or $m_k > m_k^0$. In the first case, both drifts $f_k^0(m_k)$ and $f_k^1(m_k)$ are positive, in the second case the drifts are bidirectional, while in the third case the drifts are both negative.

We have assumed that $f_k^a(\cdot)$ is non-increasing for $a = 0, 1$ and $\bar{s}_k(\bar{m}_k)$ strictly monotone in $\bar{m}_k$. In order to define the fluid index policy, we will need the following definition and the additional assumptions on the drifts.

Definition 4.1. *If $m_k^0 > m_k^1$, we set $\bar{a} = 1$ and if $m_k^1 \geq m_k^0$ we set $\bar{a} = 0$.*

Assumption 4.1. *We assume that:*

- *$f_k^a(m_k)$ is differentiable on $[m_k^{\bar{a}}, m_k^{1-\bar{a}}]$ for $a = 0, 1$.*
- *$f_k^a(m_k)$ is convex on m_k for $a = 0, 1$.*
- *$f_k^{\bar{a}}(m_k) - f_k^{\bar{a}}(\bar{m}_k) \geq (\leq) f_k^{1-\bar{a}}(m_k) - f_k^{1-\bar{a}}(\bar{m}_k)$, for all $m_k \leq (\geq) \bar{m}_k$, with $\bar{m}_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}]$.*
- *$1 - \bar{a} + (2\bar{a} - 1)\bar{s}_k(\bar{m}_k)$ is convex in $\bar{m}_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}]$.*

The hypothesis in Assumption 4.1 are easily verified for particular problems. This is done in the four examples considered in Section 4.3.

In the following proposition we give the expression for the fluid index and state an optimal solution of the fluid problem (4.1.3). The proof can be found in the Appendix and relies on the technical Lemma 4.2.

Proposition 4.1. *Let Assumption 4.1 hold. Assume $C_k(m_k, a)$, $a = 0, 1$, is differentiable for $m_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}]$, convex and non-decreasing for all m_k and $a = 0, 1$.*

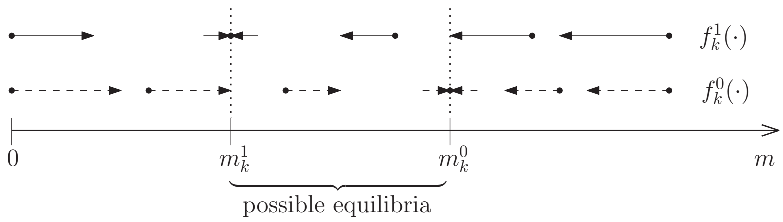


Figure 4.1: Representation of fluid equilibria and drift functions when $m_k^1 < m_k^0$.

We assume $C_k(m_k, \bar{a}) - C_k(\bar{m}_k, \bar{a}) \leq C_k(m_k, 1-\bar{a}) - C_k(\bar{m}_k, 1-\bar{a})$. *We define*

$$w_k^{(a)}(m_k) := (f_k^1(m_k) - f_k^0(m_k)) \frac{C_k(m_k, a) - C_k(m_k^a, a)}{f_k^a(m_k)}, \textit{for } a = 0, 1,$$
$$w_k^{(2)}(m_k) := \frac{(f_k^1(m_k) - f_k^0(m_k))(f_k^0(m_k)\frac{\mathrm{d}C_k(m_k,1)}{\mathrm{d}m_k} - f_k^1(m_k)\frac{\mathrm{d}C_k(m_k,0)}{\mathrm{d}m_k})}{f_k^0(m_k)\frac{\mathrm{d}f_k^1(m_k)}{\mathrm{d}m_k} - f_k^1(m_k)\frac{\mathrm{d}f_k^0(m_k)}{\mathrm{d}m_k}},$$

and assume $(2\bar{a}-1)w_k^{(i)}(m_k)$, $i = 0, 1, 2$, *is non-decreasing for all* m_k. *Then an optimal solution of* (4.1.3) *is* $s_k(t) = 1$ *if* $W \leq w_k(m_k)$ *and* $s_k(t) = 0$ *if* $W > w_k(m_k)$, *where* $w_k(m_k)$ *is a continuous function defined as follows*

$$w_k(m_k) := C_k(m_k, 0) - C_k(m_k, 1) + \begin{cases} w_k^{(\bar{a})}(m_k) & \textit{if } m_k < m_k^{\bar{a}}, \\ w_k^{(2)}(m_k) & \textit{if } m_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}], \\ w_k^{(1-\bar{a})}(m_k) & \textit{if } m_k > m_k^{1-\bar{a}}. \end{cases}$$

We will refer to the function $w_k(\cdot)$ as the *fluid index*.

The fluid index allows for the following interpretation. The function $EC_k(\bar{s}_k, W)$ is convex in $\bar{s}_k$ and $\partial EC_k(\bar{s}_k, W)/\partial \bar{s}_k = 0$ if and only if $W = C_k(m_k, 0) - C_k(m_k, 1) + w_k^{(2)}(m_k)$. That is, when $m_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}]$ the fluid index is the value of W that minimizes the cost at equilibrium m_k. This calculations can be found in the proof of Proposition 4.1. In addition, when $m_k < m_k^{\bar{a}}$ or $m_k > m_k^{1-\bar{a}}$, the fluid index $w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + w_k^{(a)}(m_k)$ can be rewritten as follows:

$$w_k(m_k) = f_k^{1-a}(m_k)\left(\frac{C_k(m_k, 0) - C_k(m_k^a, a)}{f_k^0(m_k)} - \frac{C_k(m_k, 1) - C_k(m_k^a, a)}{f_k^1(m_k)}\right),$$

for $a = 0, 1$. Note that this can be seen as the fluid equivalent of Equation (3.3.2) and provides a similar interpretation.

We observe from Proposition 4.1 that monotonicity of $w_k(m)$ in m implies that threshold policies are optimal for problem (4.1.3): in the case $m_k^1 < m_k^0$ ($m_k^1 > m_k^0$), non-decreasingness (non-increasingness) of $w_k(\cdot)$ implies that a threshold policy of structure 0-1 (1-0) is optimal, that is, it is optimal to be passive if and only if $m_k \leq m_k'(W)$ ($m_k \geq m_k'(W)$), with $m_k'(W)$ such that $w_k(m_k'(W)) = W$. This as opposed to the stochastic model, where optimality of threshold policies needs to be verified independently and might be difficult to derive.

Monotonicity of $w_k(m)$ is a simple property to verify. This represents an advantage with respect to the stochastic model, since in general optimality of threshold policies for birth-and-death stochastic bandits and indexability are difficult to establish. In Section 4.3 we will show the monotonicity of the fluid index to be satisfied for three examples. The next lemma states sufficient conditions for $w_k(\cdot)$ to be monotone.

Lemma 4.1. *Assume* $C_k(m_k,1) = C_k(m_k,0)$ *and* $\frac{\mathrm{d}f_k^1(m_k)}{\mathrm{d}m_k} = \frac{\mathrm{d}f_k^0(m_k)}{\mathrm{d}m_k}$. *Let* $C_k(m_k,1)$ *be non-decreasing in* m_k *and let* $C_k(m_k,1)$ *and* $f_k^1(m_k)$ *be polynomials of degree* $P > 0$ *and* $\alpha \geq 0$, *respectively. Then the function* $(2\bar{a}-1)w_k(m_k)$ *is non-decreasing if* $f_k^{\bar{a}}(m_k) - f_k^{1-\bar{a}}(m_k) < 0$.

Proof. The proof follows after substituting $C_k(m_k,1) = C_k(m_k,0)$ and $\frac{\mathrm{d}f_k^1(m_k)}{\mathrm{d}m_k} = \frac{\mathrm{d}f_k^0(m_k)}{\mathrm{d}m_k}$ in the expressions of Proposition 4.1, and using that $f_k^a(\cdot)$ is non-increasing for $a = 0,1$. □

In Section 2.3 we defined the indexability property that allowed us to use the index values as a heuristics for the original problem. For the fluid model we use the same definition, that is, the fluid bandit is *indexable* if the collection of states in which the optimal action is passive increases as W increases. This property follows for the fluid model directly from the fact that $D_k(W) = \{m_k : W > w_k(m_k)\}$, see Definition 2.1. This as opposed to the stochastic model, for which indexability needs to be verified independently.

Remark 4.3. *The generality of our approach is illustrated by the fact that when applied to classical problems, it retrieves well-known index policies. For instance, it can be verified that in the case of a multi-class queue with linear holding costs our fluid index becomes the optimal $c\mu$-rule, while for convex holding costs it coincides with the Generalized $c\mu$-rule (introduced and heavy-traffic optimality established in van Mieghem [70]). For a multi-class queue with user impatience and linear holding cost, our fluid index reduces to the $c\mu/\theta$-rule (introduced and asymptotic fluid optimality established in Atar et al. [8]).*

4.1.3 Fluid index policy

The property of indexability allows us to define a heuristic for (3.1.3) based on the fluid index $w_k(\cdot)$ as obtained for the fluid version of the relaxed problem.

Definition 4.2 (Fluid index policy)**.** *Assume at time t we are in state $\vec{N}(t) = \vec{n}$. The fluid index policy prescribes to serve the M bandits having currently the highest non-negative fluid index $w_k(n_k)$.*

In Section 4.3 we will present numerical simulations that show that the performance of our fluid index policy is in fact nearly optimal. In addition, we numerically compare the fluid index with Whittle's index for the stochastic model.

4.2 Equivalence of stochastic index and fluid index in light-traffic regime

In this section we consider the Whittle index as given in Corollary 2.1 and the fluid index proposed in Proposition 4.1. We first prove that they become equivalent in a light-traffic regime. Second, we prove that for large values of states Whittle's index and fluid index coincide for the multi-class abandonment model considered in Chapter 3.

Recall that the birth transition rates are given by $b_k^a(\cdot)$ for $a = 0,1$. Let us first assume that $b_k^a(m_k) = \lambda\gamma_k$ for all m_k, in problems where 0-1 type of threshold policies are optimal, and $b_k^a(m_k) = \lambda\gamma_k a$ for all m_k, in problems where 1-0 type of threshold policies are optimal. Here $\sum_{k=1}^{K}\gamma_k = 1$, and λ represents the intensity of the input to the system. We consider the light-traffic regime $\lambda \to 0$, that is, the birth rate is close to 0.

We first compute the Whittle index in light traffic in Proposition 4.2. The proof can be found in Appendix 4.4.3.

Proposition 4.2. *Let $W_k(\cdot)$ be as given in* (2.3.6). *Then $W_k(n_k) = W_k^{LT}(n_k) + o(1)$ as $\lambda \downarrow 0$, where*

$$W_k^{LT}(n_k) := C_k(n_k, 0) - C_k(n_k, 1) + (d_k^1(n_k) - d_k^0(n_k))\frac{C_k(n_k, \bar{a}) - C_k(0, \bar{a})}{d_k^{\bar{a}}(n_k)},$$

and $\bar{a} = 0$ if the threshold that solves (2.3.3) *is of 0-1 type and $\bar{a} = 1$ if the threshold that solves* (2.3.3) *is of 1-0 type.*

By assumption $d_k^a(m_k) = \lambda\gamma_k > 0$ for all $m_k > 0$ and $a = 0, 1$, which implies $m_k^0 \to 0$ and $m_k^1 \to 0$ as $\lambda \to 0$. Hence, the fluid index in the light-traffic regime is given by $w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + w_k^{(\bar{a})}(m_k)$. In the next proposition we establish the equivalence of the Whittle index and the fluid index in the light-traffic regime. The proof can be found in Appendix 4.4.4.

Proposition 4.3. *Assume $d_k^a(m_k) > 0$ for all $m_k > 0$ and $a = 0, 1$. Let $W_k(\cdot)$ be given as in* (2.3.6). *Assume an optimal solution for* (4.1.3) *to be the threshold policy n_k. Then*

$$\lim_{\lambda \downarrow 0} W_k(m_k) = \lim_{\lambda \downarrow 0} w_k(m_k).$$

Let us now consider the multi-class abandonment model considered in Chapter 3, that is, $b_k^a(m_k) = \lambda_k$ and $d_k^a(m_k) := (\mu_k + \theta_k' - \theta_k)a + \theta_k m_k$. In the next proposition we show that the fluid index $w_k(n_k)$ coincides with Whittle's index as given in (2.3.6) for large values of the state, when the cost functions $C_k(m, 1)$ and $C_k(m, 0)$ for all m, are upper bounded by polynomial functions of finite degrees $P_k < \infty$ and $Q_k < \infty$, respectively. Hence, we can write $C_k(n_k, a) = E_k(n_k, a) + o(1)$, for large values of n_k, where $E_k(n_k, 1) = \sum_{i=0}^{P_k} C_k^{(P_k,i)} n_k^i$, with

$$C_k^{(P_k,i)} := \lim_{n_k \to \infty} \frac{C_k(n_k, 1) - \sum_{j=i+1}^{P_k} C_k^{(P_k,j)} n_k^j}{n_k^i},$$

and $E_k(n_k, 0) = \sum_{i=0}^{Q_k} E_k^{(Q_k,i)} n_k^i$, with

$$E_k^{(Q_k,i)} := \lim_{n_k \to \infty} \frac{C_k(n_k, 0) - \sum_{j=i+1}^{Q_k} E_k^{(Q_k,j)} n_k^j}{n_k^i}.$$

Proposition 4.4. *Assume that $C_k(n_k, 1)$ and $C_k(n_k, 0)$ are upper bounded by a polynomial of degree P_k and Q_k, respectively, with $Q_k > P_k$. Then, for the abandonment problem we have*

$$\lim_{n_k \to \infty} \frac{W_k(n_k)}{w_k(n_k)} = 1. \tag{4.2.1}$$

If we further assume $P_k = Q_k$ and $C_k^{(P_k,i)} = E_k^{(P_k,i)}$ for all $i \in \{2, \ldots, P_k\}$, then as $n_k \to \infty$,

$$W_k(n_k) = w_k(n_k) + o(1). \tag{4.2.2}$$

As an example we consider $C_k(n_k, a) = C_k(n_k)$ or $C_k(n_k, a) = C_k((n_k - a)^+)$. Then $Q_k = P_k$, and hence (4.2.1) holds. In case, $C_k(n_k, a) = C_k(n_k)$, then in addition (4.2.2) holds. The proof of Proposition 4.4 can be found in Appendix 4.4.5.

4.3 Case studies

In this section we evaluate both the stochastic and fluid index policies for four examples of birth-and-death bandits. The objective is to show how these policies apply to the following four decision making problems: (i) scheduling in a multi-class single-server abandonment queue, (ii) opportunistic scheduling in a wireless downlink channel, which both belong to the class of problems depicted in Figure 2.1 (left), (iii) optimal blocking/routing in a power-aware server farm, and (vi) scheduling a make-to-stock queue with perishable items, which both belong to the class of problems depicted in Figure 2.1 (right). In all cases, we compare the performance of Whittle's index policy (2.3.6) and the fluid index policy, as given in Proposition 4.1, against the optimal policy, which is computed using the value iteration approach, see Puterman [77, Chapter 8.5] and also Section 1.3.3 where this technique was introduced. Our overall conclusion is that the performance of the Whittle and the fluid index policies is nearly optimal.

4.3.1 Scheduling in a multi-class abandonment queue

In this section we consider the multi-class abandonment queue considered in Chapter 3. Class-k users arrive according to a Poisson process of rate λ_k and their service requirement is exponentially distributed with mean $1/\mu_k$. Customers waiting in the queue abandon after an exponentially distributed amount of time with mean $1/\theta_k$, and customers that are already receiving service abandon the system after an exponentially distributed amount of time with mean $1/\theta'_k$. The transition rates of this MDP write therefore $b_k^a(n_k) = \lambda_k$, and $d_k^a(n_k) = (\mu_k + \theta'_k)a + \theta_k(n_k - a)^+$ for $a = 0, 1$. And recall the assumption $\mu_k + \theta'_k \geq \theta_k$.

The objective is to minimize the average holding cost, where as defined in Equation (3.1.4), $\tilde{C}_k(N_k, a)$ is the sum of the holding cost and abandonment cost when having N_k class-k users in the system.

In this setting an optimal policy of the relaxed optimization problem is of threshold type with 0-1 structure. The latter was proven in Proposition 3.1.

The steady-state probabilities (obtained using the standard formula for a birth-and-death process) of class k under threshold policy n_k are given by Equation (3.2.3). These expressions were used in Proposition 3.2 to prove indexability of the model under consideration. Hence, from Corollary 2.1 we have that Whittle's index is given by (2.3.6) in case (2.3.6) is non-decreasing. In Chapter 3 we have derived many properties of Whittle's index (2.3.6) for this model. We are now interested in deriving the fluid index, which has a tractable closed-form expression for any cost function.

The fluid dynamics is given by

$$\begin{aligned}\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} =& \lambda_k - s_k(t)(\mu_k + \theta'_k + \theta_k(m_k(t) - 1)) - (1 - s_k(t))\theta_k m_k(t)\\ =& \lambda_k - (\mu_k + \theta'_k - \theta_k)s_k(t) - \theta_k m_k(t),\end{aligned}$$

where $s_k(t) \in \{0, 1\}$. Hence, $m_k^0 = \lambda_k/\theta_k$ and $m_k^1 = (\lambda_k - (\mu_k + \theta'_k - \theta_k))/\theta_k$, that is, the equilibrium points $\bar{m}_k$ satisfy $\bar{m}_k \in [\max(0, m_k^1), \lambda_k/\theta_k]$. From Proposition 4.1 we can now derive the fluid index, which describes the policy that minimizes the bias-optimal criteria as given in (4.1.3). We first derive the fluid index and we then prove that the fluid index policy is optimal, see Proposition 4.5 and Proposition 4.6, respectively.

Proposition 4.5. *Assume $C_k(m_k, a)$ is differentiable, convex and non-decreasing in m_k. In addition, assume $C_k(m_k, 0) - C_k(m_k, 1)$ to be convex non-decreasing in m_k. Then, the fluid index is non-decreasing and given by:*

$$
\begin{aligned}
w_k(m_k) :=& C_k(m_k,0) - C_k(m_k,1) + \delta_k(\mu_k + \theta'_k) - \delta'_k\theta'_k \\
&+ \begin{cases} w_k^{(1)}(m_k) & \text{if } 0 \le m_k < \max\left(0, \frac{\lambda_k - (\mu_k + \theta'_k - \theta_k)}{\theta_k}\right), \\ w_k^{(2)}(m_k) & \text{if } \max\left(0, \frac{\lambda_k - (\mu_k + \theta'_k - \theta_k)}{\theta_k}\right) \le m_k \le \frac{\lambda_k}{\theta_k}, \\ w_k^{(0)}(m_k) & \text{if } m_k > \frac{\lambda_k}{\theta_k}, \end{cases}
\end{aligned} \tag{4.3.1}
$$

where

$$
\begin{aligned}
w_k^{(1)}(m_k) =& \frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k} \frac{\left(C\left(\frac{\lambda_k - (\mu_k + \theta'_k - \theta_k)}{\theta_k}, 1\right) - C(m_k, 1)\right)}{(\lambda_k - (\mu_k + \theta'_k - \theta_k))/\theta_k - m_k}, \\
w_k^{(2)}(m_k) =& \frac{(\lambda_k - \theta_k m_k)\frac{\mathrm{d}}{\mathrm{d}m_k}C_k(m_k,1) + (\theta_k m_k + \mu_k + \theta'_k - \theta_k - \lambda_k)\frac{\mathrm{d}}{\mathrm{d}m_k}C_k(m_k,0)}{\theta_k}, \\
w_k^{(0)}(m_k) =& \frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k} \frac{\left(C_k(m_k,0) - C_k\left(\frac{\lambda_k}{\theta_k}, 0\right)\right)}{m_k - \lambda_k/\theta_k}.
\end{aligned}
$$

Proof. Equation (4.3.1) being non-decreasing follows from observing that for any convex non-decreasing function $C_k(m_k, 1)$, for $m_k \le m'_k$, the function $\frac{C_k(m'_k,a) - C_k(m_k,a)}{m'_k - m_k}$, is non-decreasing in m_k for $a = 0, 1$, and the fact that $C_k(m_k, 0) - C_k(m_k, 1)$ being non-decreasing implies $\frac{\mathrm{d}_k C_k(m_k,0)}{\mathrm{d}m_k} \ge \frac{\mathrm{d}_k C_k(m_k,1)}{\mathrm{d}m_k}$. Equation (4.3.1) now follows from Proposition 4.1 together with Lemma 4.6. ☐

In the following proposition we present a bias-optimal solution for the fluid problem (4.1.3).

Proposition 4.6. *Assume the same conditions as in Proposition 4.5. An optimal solution for problem* (4.1.3) *with transitions rates* (3.2.3) *is: $s_k(t) = 1$ if $W \le w_k(m_k)$ and $s_k(t) = 0$ if $W > w_k(m_k)$, where $w_k(m_k)$ is as defined in Proposition 4.5.*

Proof. The proof follows by verifying Assumption 4.1. We have $f_k^a(m_k) = \lambda_k - (\mu_k + \theta'_k - \theta_k)a - \theta_k m_k$, for $a \in \{0, 1\}$. Differentiability of $f_k^a(m_k)$ follows directly, also monotonicity of $\bar{s}_k(\bar{m}_k)$ for $\bar{m}_k$ in $[m_k^1, m_k^0]$ and convexity of $f_k^0(m_k)$. The function $f_k^1(m_k)$ satisfies

$$\frac{\mathrm{d}^2 f_k^1(m_k)}{\mathrm{d}m_k^2} = 0,$$

for all $m_k \ge 0$, and it is hence convex in m_k.

We have, $\bar{s}_k(\bar{m}_k) = (\lambda_k - \theta_k m_k)/(\mu_k + \theta'_k - \theta_k)$, hence

$$\frac{\mathrm{d}^2}{\mathrm{d}m_k^2}\left(\bar{s}_k(\bar{m}_k)\right) = 0,$$

that is, $\bar{s}_k(\bar{m}_k)$ is convex for $\bar{m}_k \in [m_k^1, m_k^0]$.

The inequality $f_k^1(m_k) - f_k^1(\bar{m}_k) \geq (\leq) f_k^0(m_k) - f_k^0(\bar{m}_k)$ for all $m_k \leq (\geq)\bar{m}_k$ and $\bar{m}_k \in [m_k^1, m_k^0]$, simplifies to $\theta_k(\bar{m}_k - m_k - \bar{m}_k + m_k) \geq (\leq) 0$ for all $m_k \leq (\geq)\bar{m}_k$ and $\bar{m}_k \in [m_k^1, m_k^0]$, which is always satisfied.

Then from Propositions 4.1 and 4.7, an optimal solution for problem (4.1.3) is $s_k(t) = 1$ if $W \leq w_k(m_k)$ and $s_k(t) = 0$ if $W > w_k(m_k)$. □

We directly observe that for linear holding cost, the fluid index is state-independent and coincides with that of the stochastic model as stated in Proposition 3.3. Now assume $C_k(m_k, a_k) = C_k(m_k)$, that is, holding cost for customers in the system. In that case, the fluid index simplifies as follows:

$$w_k^{(2)}(m_k) = \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k} \frac{\mathrm{d}}{\mathrm{d}m_k} C_k(m_k),$$

which corresponds to the $C'(m)\mu/\theta$-rule when $\theta_k' = \theta_k$. We refer to this rule as the Generalized $c\mu/\theta$-rule ($Gc\mu/\theta$). The terms $w_k^{(1)}(m_k)$ and $w_k^{(0)}(m_k)$ reduce to

$$\begin{aligned} w_k^{(1)}(m_k) &= \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k} \frac{(C_k((\lambda_k - (\mu_k + \theta_k' - \theta_k))/\theta_k) - C_k(m_k))}{(\lambda_k - (\mu_k + \theta_k' - \theta_k))/\theta_k - m_k}, \\ w_k^{(0)}(m_k) &= \frac{(\mu_k + \theta_k' - \theta_k)}{\theta_k} \frac{(C_k(m_k) - C_k(\lambda_k/\theta_k))}{m_k - \lambda_k/\theta_k}. \end{aligned}$$

We refer to Bispo [26] where index policies based on first-order difference have also been proposed and are shown to empty the system with the lowest cost possible in a single server multi-class queue without abandonments and no future arrivals.

The numerical evaluation of both the Whittle index policy and the fluid index policy for the abandonment model have been presented in Chapter 3, we therefore refer the reader to Chapter 3 (Section 3.6) for a discussion on the numerical performance and near-optimality of these heuristics.

4.3.2 Opportunistic scheduling in a wireless downlink channel

In this section we consider a wireless downlink channel shared by K classes of users. Class-k users arrive according to a Poisson process of rate λ_k and their service requirement is exponentially distributed with mean $1/\tilde{\mu}_k$. At any moment in time, the base station can send data to at most one of the users present in the system. We assume the channel quality of a class-k user to be independent of the other users and can be modeled with a uniform random variable G_k on $[0, \gamma_k)$. As a consequence of opportunistic scheduling, the capacity when serving class k is the maximum of N_k i.i.d. random variables $G_{k,1}, \ldots, G_{k,N_k}$, distributed as G_k, see Borst [28]. Hence, the expected capacity is given by $\mathbb{E}(\max(G_{k,1}, \ldots, G_{k,N_k})) = \gamma_k N_k(t)/(N_k(t) + 1)$. We therefore take as departure rate $\mu_k(m_k) = \mu_k m_k/(m_k + 1)$, where $\mu_k := \tilde{\mu}_k \gamma_k$. This Markov decision process is characterized by the following transition rates:

$$b_k^a(m_k) = \lambda_k, \text{ and } d_k^a(m_k) = \mu_k \frac{m_k}{m_k + 1} a, \tag{4.3.2}$$

where $a = 1$ ($a = 0$) stands for serving (not serving) class k, see Figure 2.1. In order for the system to be stable we assume $\rho := \sum_{k=1}^K \lambda_k/\mu_k < 1$.

The objective is to minimize the average holding cost, where $C_k(m_k, a)$ is the holding cost when having m_k class-k users in the system. Note that $C_k(m_k, a) = C_k(m_k)$ represents holding costs for users in the *system*, while $C_k(m_k, a) = C_k((m_k - a)^+)$ represents holding costs for users in the *queue*.

In this setting an optimal policy of problem (2.3.3) is of threshold type with 0-1 structure. The latter follows directly from Proposition 2.1.

The steady-state probabilities (obtained using the standard formula for a birth-and-death process) of class k under threshold policy n_k are given by

$$
\begin{aligned}
&\pi_k^{n_k}(m_k) = 0, \forall m_k \leq n_k - 1,\\
&\pi_k^{n_k}(m_k) = \left(\frac{\lambda_k}{\mu_k}\right)^{m_k - n_k} \frac{m_k + 1}{n_k + 1} \pi^{n_k}(n_k), \forall m_k \geq n_k + 1,\\
&\pi_k^{n_k}(n_k) = 1/\left(1 + \frac{1}{n_k + 1} \sum_{i=1}^{\infty} \left(\frac{\lambda_k}{\mu_k}\right)^i (n_k + 1 + i)\right).
\end{aligned}
$$

We now check that the function $\sum_{i=0}^{n} \pi_k^n(i)$ is strictly increasing in n, that is, $\pi_k^n(n) \leq \pi_k^{n+1}(n+1)$. To do so observe that after some algebra $\pi_k^n(n) \leq \pi_k^{n+1}(n+1)$ simplifies to $(\frac{1}{n+1} - \frac{1}{n+2}) \sum_{i=1}^{\infty} (\frac{\lambda_k}{\mu_k})^i, i \geq 0$, which is satisfied for all n. From Proposition 2.2 (and consequently Corollary 2.1) we have now that the problem is indexable and that Whittle's index is given by (2.3.6) in case (2.3.6) is non-decreasing. Equation (2.3.6) can be numerically computed and verified to be non-decreasing. However, it does not help to provide insights into the properties of Whittle's index policy. This is the main motivation to derive the fluid index, which has a tractable closed-form expression.

The fluid dynamics is given by $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \lambda_k - \mu_k \frac{m_k}{m_k+1} s_k(t)$, where $s_k(t) \in \{0, 1\}$ ($s_k(t) = 1$ if station k is activated). Hence, $m_k^0 = \infty$ and $m_k^1 = \lambda_k/(\mu_k - \lambda_k)$, that is, the equilibrium points satisfy $\bar{m}_k \in [m_k^1, \infty)$. From Proposition 4.1 we can now derive the fluid index, which describes the policy that minimizes the bias-optimal criteria as given in (4.1.3).

In the following proposition we derive the fluid index.

Proposition 4.7. *Assume $C_k(m_k, a)$ is differentiable, convex and non-decreasing in m_k. In addition, assume $C_k(m_k, 0) - C_k(m_k, 1)$ to be convex non-decreasing in m_k. Then, the fluid index is non-decreasing and given by:*

$$
w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + \begin{cases} w_k^{(1)}(m_k) & \text{if } m_k < \lambda_k/(\mu_k - \lambda_k),\\ w_k^{(2)}(m_k) & \text{if } \lambda_k/(\mu_k - \lambda_k) \leq m_k, \end{cases} \tag{4.3.3}
$$

where

$$
\begin{aligned}
w_k^{(1)}(m_k) &= \mu_k m_k \frac{C_k((\lambda_k/(\mu_k - \lambda_k), 1) - C_k(m_k, 1)}{\lambda_k - (\mu_k - \lambda_k) m_k},\\
w_k^{(2)}(m_k) &= m_k(m_k + 1) \left(\frac{\mathrm{d}C_k(m_k, 1)}{\mathrm{d}m_k} - \frac{\mathrm{d}C_k(m_k, 0)}{\mathrm{d}m_k}\right) + \frac{m_k^2 \mu_k}{\lambda_k} \frac{\mathrm{d}C_k(m_k, 0)}{\mathrm{d}m_k}.
\end{aligned}
$$

Proof. Equation (4.3.3) being non-decreasing follows from observing that for any convex non-decreasing function $C_k(m_k, 1)$, for $m_k \leq m_k'$, the function $\frac{C_k(m_k', 1) - C_k(m_k, 1)}{m_k' - m_k}$, is non-decreasing in m_k and from the fact that $C_k(m_k,0) - C_k(m_k,1)$ is convex and non-decreasing implies that $\frac{\mathrm{d}_k C_k(m_k,0)}{\mathrm{d}m_k} - \frac{\mathrm{d}_k C_k(m_k,1)}{\mathrm{d}m_k} \geq 0$ and is non-decreasing. Equation (4.3.3) now follows from Proposition 4.1 together with Lemma 4.8. □

In the following proposition we present a bias-optimal solution for the fluid problem (4.1.3).

Proposition 4.8. *Assume $C_k(m_k,a)$ is differentiable, convex and non-decreasing in m_k. In addition, assume $C_k(m_k,0)-C_k(m_k,1)$ to be convex non-decreasing in m_k. An optimal solution for problem* (4.1.3) *with transitions rates* (4.3.2) *is: $s_k(t)=1$ if $W \leq w_k(m_k)$ and $s_k(t)=0$ if $W > w_k(m_k)$, where $w_k(m_k)$ is as defined in Proposition 4.7.*

Proof. The proof follows by verifying Assumption 4.1. We have $f_k^a(m_k) = \lambda_k - \mu_k \frac{m_k}{m_k+1} a$, for $a \in \{0,1\}$. Differentiability of $f_k^a(m_k)$ follows directly, also monotonicity of $\bar{s}_k(\bar{m}_k)$ for $\bar{m}_k$ in $[m_k^1, m_k^0]$ and convexity of $f_k^0(m_k)$. The function $f_k^1(m_k)$ satisfies

$$\frac{\mathrm{d}^2 f_k^1(m_k)}{\mathrm{d}m_k^2} = \mu_k \frac{2}{(m_k+1)^3} \geq 0,$$

for all $m_k \geq 0$, and it is hence convex in m_k.

We have, $\bar{s}_k(\bar{m}_k) = \lambda_k(\bar{m}_k+1)/\mu_k \bar{m}_k$, hence

$$\frac{\mathrm{d}^2}{\mathrm{d}m_k^2}\left(\bar{s}_k(\bar{m}_k)\right) = 2\frac{\lambda_k}{\mu_k \bar{m}_k^3} \geq 0,$$

that is, $\bar{s}_k(\bar{m}_k)$ is convex for $\bar{m}_k \in [m_k^1, m_k^0]$.

The inequality $f_k^1(m_k) - f_k^1(\bar{m}_k) \geq (\leq) f_k^0(m_k) - f_k^0(\bar{m}_k)$ for all $m_k \leq (\geq) \bar{m}_k$ and $\bar{m}_k \in [m_k^1, m_k^0]$, simplifies to $\mu_k(\frac{\bar{m}_k}{\bar{m}_k+1} - \frac{m_k}{m_k+1}) \geq (\leq) 0$ for all $m_k \leq (\geq) \bar{m}_k$ and $\bar{m}_k \in [m_k^1, m_k^0]$, which is satisfied due to $\frac{m_k}{m_k+1}$ being a non-decreasing function in m_k.

Then from Proposition 4.1 and 4.7, an optimal solution for problem (4.1.3) is $s_k(t)=1$ if $W \leq w_k(m_k)$ and $s_k(t)=0$ if $W > w_k(m_k)$. □

The fluid index being non-decreasing implies that the fluid index policy as defined in Section 4.1.3 will give more importance to a class to be served as its queue length grows.

Having a closed-form expression for the fluid index as given in (4.3.3) gives us insights on the behavior of the system with respect to the parameters involved. For the sake of clarity assume linear cost of type $C_k(m,0) = C_k(m,1) = c_k m$, and $\lambda_k = \lambda \delta_k$. Then $w_k(m_k) = c_k m_k \mu_k/(\mu_k - \lambda_k)$ for $m_k < \lambda_k/(\mu_k - \lambda_k)$, and $w_k(m_k) = c_k m_k^2 \mu_k/\lambda_k$ otherwise. Hence, as $\lambda \downarrow 0$, in states very close to the origin priority is given according to $c_k m_k \mu_k/(\mu_k - \lambda_k)$ and otherwise according to $c_k \mu_k m_k^2/\delta_k$.

In the example below we will numerically evaluate the performance of the two index policies against the optimal policy.

Example 1. Let us assume two classes of users with $\mu_1 = 16, \mu_2 = 27$, and $\lambda_1/\mu_1 = \rho/2$, $\lambda_2/\mu_2 = \rho/2$. We further assume that the cost function is given by $C_k(n,a) = c_k(n-a)^2 + b_k(n-a)$ for $k \in \{1,2\}$, with $b_1 = 0.1, b_2 = 1, c_1 = 2$ and $c_2 = 1.5$, that is, quadratic holding cost for the number of users waiting to be served. We compute the relative error of the Whittle index policy as well as the relative error of the fluid index policy with respect to the optimal policy, see Table 4.1. We observe that both index policies perform nearly optimal across all loads.

Table 4.1: Example 1: Relative sub optimality gap in %.

ρ	0.1	0.2	0.3	0.4
Whittle index policy	0.20289	1.16215	2.54794	3.54934
Fluid index policy	0.20289	1.16215	2.55440	3.54936
ρ	0.5	0.6	0.7	0.8
Whittle index policy	3.52057	2.54793	1.56715	0.66077
Fluid index policy	3.52098	2.55439	1.60799	0.75140

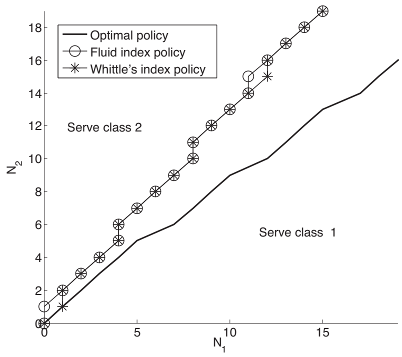


Figure 4.2: Switching curves under the optimal policy, Whittle's index policy and fluid index policy.

In Figure 4.2 we depict the actions taken under the optimal policy, Whittle's index policy, and the fluid index policy, for $\rho = 0.5$. The three policies are characterized by the three switching curves as depicted in the figure. Below the curve class 1 is served and above the curve class 2 is served. We observe that the two switching curves corresponding to the fluid index policy and the Whittle index policy coincide in almost the entire state space, and capture the qualitative structure of the optimal policy.

4.3.3 Routing/blocking in a power-aware server-farm

We consider a server farm with K heterogeneous service stations each having one server, see Figure 2.1 (right). Users arrive to the system following a Poisson process of rate λ. An arriving user is either routed to one of the stations or is blocked. The service capacity of the power-aware servers follows a speed-scaling rule Wierman [100] in order to balance between power consumption and server capacity. We assume that when in state m_k, the service capacity is $c(m_k) := m_k^\alpha$, with $0 < \alpha < 1$. The service requirement of a user in station k is exponentially distributed with mean $1/\mu_k$. Hence, the departure rate is $\mu_k(m_k) = \mu_k m_k^\alpha$.

Each time a user is blocked for service a penalty D is paid, hence, implying blocking cost to occur at rate λD. A common model for the power consumption is $c(m_k)^{1/\alpha}$, hence, we have that the power consumed in state m_k is equal to m_k. We therefore take as cost $C_k(m_k, a) = C_k(m_k) + \beta_k m_k + D\lambda(1 - a)$, where $C_k(m_k)$ represents the holding cost of having users in server k and $\beta_k \geq 0$ controls the relative cost of power consumption. We assume $C_k(m_k)$ to be convex. The aim is to find the optimal blocking/routing policy in order to minimize the sum of the average holding cost, power consumption and penalty for blocking users. An optimal load balancing policy must strike the right balance between dispatching a user to a server with a large queue length (which implies a large increase in holding cost, due to convexity, but a high service rate), dispatching to a server with a small queue length (which implies a small increase in holding cost but a small service rate), and blocking a user (which implies a blocking cost, however no additional holding cost is incurred). This is a very complex optimization problem. We will see that the two index policies as described in this paper are able to perform close to optimal.

The Markov chain has the following transitions:

$$b_k^a(m_k) = \lambda a, \text{ and } d_k^a(m_k) = \mu_k m_k^\alpha, \tag{4.3.4}$$

where $a = 0$ ($a = 1$) stands for blocking (accepting) a user in server k.

For the model under study, 1-0 type of threshold policies are an optimal solution of problem (2.3.3). The proof follows from Proposition 2.1.

The steady-state probabilities (obtained using the standard formula for a birth-and-death process) of class k under threshold policy n_k are given by

$$\begin{aligned}
\pi_k^{n_k}(m_k) &= 0, \forall m_k \geq n_k + 2, \\
\pi_k^{n_k}(m_k) &= \frac{\lambda^{m_k}}{\mu_k^{m_k}(m_k!)^\alpha}\pi^{n_k}(0), \forall m_k \leq n_k + 1, \\
\pi_k^{n_k}(0) &= \left(\sum_{m_k=0}^{n_k+1} \frac{\lambda^{m_k}}{\mu_k^{m_k}(m_k!)^\alpha}\right)^{-1}.
\end{aligned}$$

We now check that the function $\sum_{i=0}^{n} \pi_k^n(i)$ is strictly increasing in n, or equivalently, $\pi_k^n(n+1)$ is strictly decreasing in n. We then want to prove $\pi_k^n(n+1) \leq \pi_k^{n-1}(n)$ which after some algebra simplifies to

$$\begin{aligned}
\pi^{n-1}(0) \geq \frac{\lambda}{\mu(n+1)^\alpha}\pi^n(0) &\Leftrightarrow \sum_{m_k=0}^{n+1} \frac{\lambda^{m_k}}{\mu_k^{m_k}(m_k!)^\alpha} \geq \frac{\lambda}{\mu(n+1)^\alpha} \sum_{m_k=0}^{n} \frac{\lambda^{m_k}}{\mu_k^{m_k}(m_k!)^\alpha} \\
&\Leftrightarrow \sum_{m_k=1}^{n} \frac{\lambda^{m_k}}{\mu_k^{m_k}((m_k-1)!)^\alpha}\left(\frac{1}{m_k^\alpha} - \frac{1}{(n+1)^\alpha}\right) + 1 \geq 0,
\end{aligned}$$

the last inequality is satisfied due to $1/m_k^\alpha - 1/(n+1)^\alpha \geq 0$ for all $m_k \in \{1, \ldots, n\}$. From Proposition 2.2 (and consequently Corollary 2.1) we have now that the problem is indexable and that Whittle's index is given by (2.3.6) in case (2.3.6) is non-decreasing. Equation (2.3.6) can be numerically computed and verified to be non-decreasing. Note that the optimal structure of the fluid version of the relaxed optimization problem is also of 1-0 structure (since the fluid index is non-increasing).

We first determine the fluid index policy for this model. The fluid dynamics is given by $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \lambda s_k(t) - \mu_k m_k^\alpha$, with $s_k(t) \in \{0, 1\}$. We have $m_k^0 = 0$, and $m_k^1 = (\lambda/\mu_k)^{1/\alpha}$, that is, the equilibrium points are in the interval $\bar{m}_k \in [0, m_k^1]$.

In the following proposition we derive the fluid index. The proof follows from Proposition 4.1 and Lemma 4.1.

Proposition 4.9. *Assume $C_k(m_k)$ to be a polynomial of degree P with $P > \alpha$. Then, the fluid index is non-increasing and given by:*

$$w_k(m_k) = D\lambda + \begin{cases} w_k^{(2)}(m_k) & \text{if } 0 \le m_k \le (\lambda/\mu_k)^{\alpha^{-1}}, \\ w_k^{(1)}(m_k) & \text{if } (\lambda/\mu_k)^{\alpha^{-1}} < m_k, \end{cases}$$

where

$$w_k^{(2)}(m_k) = -\frac{\lambda\alpha^{-1}m_k}{\mu_k m_k^{\alpha}}\frac{\mathrm{d}\tilde{C}_k(m_k)}{\mathrm{d}m_k}$$
$$w_k^{(1)}(m_k) = -\lambda\frac{(\tilde{C}_k((\lambda/\mu_k)^{\alpha^{-1}}) - \tilde{C}_k(m_k))}{\lambda - \mu_k m_k^{\alpha}},$$

with $\tilde{C}_k(m_k) = C_k(m_k) + \beta_k m_k$.

In the following proposition we present an optimal solution for problem (4.1.3).

Proposition 4.10. *Assume $C_k(m_k)$ to be a polynomial of degree P with $P > \alpha$. An optimal solution for problem* (4.1.3) *with transition rates* (4.3.4) *is: $s_k(t) = 1$ if $W \le w_k(m_k)$ and $s_k(t) = 0$ if $W > w_k(m_k)$, where $w_k(m_k)$ is as given in Proposition 4.9.*

Proof. The proof follows by verifying Assumption 4.1. We have $f_k^a(m_k) = \lambda a - \mu_k m_k^{\alpha}$, for $a \in \{0, 1\}$ and $\alpha < 1$. Differentiability of $f_k^a(m_k)$ follows directly as well as monotonicity of $\bar{s}_k(\bar{m}_k)$ in $[m_k^0, m_k^1]$. The function $f_k^a(m_k)$ satisfies

$$\frac{\mathrm{d}^2 f_k^a(m_k)}{\mathrm{d}m_k^2} = -\alpha(\alpha - 1)\mu_k m_k^{\alpha-2} \ge 0,$$

for all $m_k \ge 0$, due to the assumption $\alpha < 1$. Hence, $f_k^a(m_k)$ is convex in m_k for $a = 0, 1$.

We have $1 - \bar{s}_k(\bar{m}_k) = 1 - \mu_k m_k^{\alpha}/\lambda$, hence

$$\frac{\mathrm{d}^2}{\mathrm{d}m_k^2}(1 - \bar{s}_k(\bar{m}_k)) = -\frac{\mu_k\alpha(\alpha - 1)\bar{m}_k^{\alpha-2}}{\lambda} \ge 0,$$

for all $\bar{m}_k \in [m_k^0, m_k^1]$, since $\alpha < 1$. Hence, $1 - \bar{s}_k(\bar{m}_k)$ is convex in $\bar{m}_k$.

The inequality $f_k^1(m_k) - f_k^1(\bar{m}_k) \le (\ge) f_k^0(m_k) - f_k^0(\bar{m}_k)$ simplifies to $\mu_k(\bar{m}_k^{\alpha} - m_k^{\alpha}) \le (\ge) \mu_k(\bar{m}_k^{\alpha} - m_k^{\alpha})$, which is satisfied for all m_k and $\bar{m}_k$.

Then from Proposition 4.1 and 4.9, an optimal solution for problem (4.1.3) is $s_k(t) = 1$ if $W \le w_k(m_k)$ and $s_k(t) = 0$ if $W > w_k(m_k)$. □

The fluid index being non-increasing implies that the fluid index policy will prefer to route to servers having a relatively small queue length. Since the fluid index policy only routes to servers with a positive fluid index, there is an $\overline{N}_k$ such that when $N_k \ge \overline{N}_k$, no users will be routed to this server k.

As in the previous section we use Proposition 4.9 to obtain interesting insights for particular cases. For the sake of clarity assume linear cost of type $C_k(m) = c_k m$. Then, as $\lambda \uparrow \infty$, $w_k(m_k)$ will be given by $D\lambda + w_k^{(2)}(m_k)$, and $w_k^{(2)}(m_k) = -\lambda c_k \frac{m_k^{1-\alpha}}{\mu_k \alpha}$, hence priority will be given according to $c_k \frac{m_k^{1-\alpha}}{\mu_k \alpha}$.

We now present an example to evaluate the performance of both index policies.

Table 4.2: Example 2: Relative sub optimality gap in %.

ρ	0.1	0.3	0.5
Fluid index	0.08704×10^{-7}	0.16036×10^{-7}	0.13968×10^{-7}
Whittle's index	0.08704×10^{-7}	0.16036×10^{-7}	0.13968×10^{-7}
ρ	0.7	0.9	1.1
Fluid index	$0.06279\ \times10^{-7}$	$0.08210\ \times10^{-7}$	$0.06124\ \times10^{-7}$
Whittle's index	0.06279×10^{-7}	0.08210×10^{-7}	0.06124×10^{-7}
ρ	1.5	2	2.5
Fluid index	0.01872×10^{-7}	0.06099×10^{-7}	0.10921×10^{-7}
Whittle's index	0.01872×10^{-7}	0.06099×10^{-7}	0.07110×10^{-7}

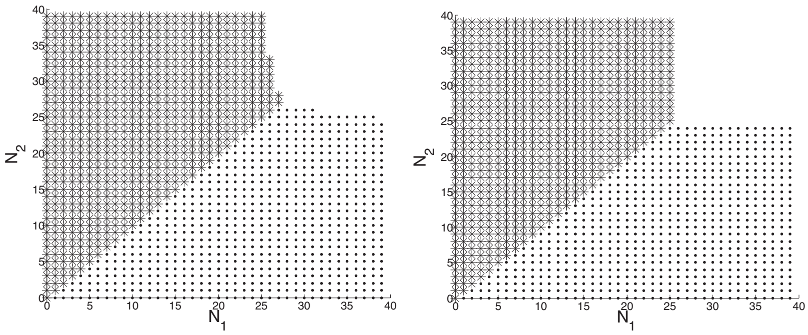


Figure 4.3: In the area with "*" (".") class 1 (class 2) is prioritized and in the white area users are blocked. Left: Optimal policy. Right: Whittle index policy and the fluid index policy.

Example 2. In this example we assume 2 classes of users which arrive at rate $\lambda = 18$. We set the speed scaling parameter at $\alpha = 1/2$. The cost function is such that $C_k(m_k, a) = C_k(m_k) + \beta_k m_k + D\lambda a$, and we assume $C_k(m_k) = c_k m_k^2$ where $c_1 = c_2 = 2$, and $\beta_1 = 3, \beta_2 = 5$. We further assume that the cost for blocking users is $D = 25$. The service rates μ_1, μ_2 are such that $\mu_1 = \mu_2 = 2\lambda/\rho$. We set $M = 1$, that is, a customer can be routed to at most one server. We observe in Table 4.2 that the performance for the Whittle index policy as well as for the fluid index policy for various values of ρ is nearly optimal. Moreover, in Figure 4.3 we illustrate the optimal strategy together with the Whittle index policy for $\rho = 2.5$. The fluid index policy coincides with the strategy given by Whittle's index policy and captures the qualitative structure of the optimal policy.

4.3.4 Scheduling a make-to-stock queue with perishable items

We consider a single production machine which produces K different classes of items. This problem belongs to the class of problems depicted in Figure 2.1 (right). Demands for a class-k item arrive to the system following a Poisson process of rate λ_k. The machine can only produce one item at a time, and the production time is exponentially distributed with mean $1/\mu_k$. The items, once produced, are stocked until either a user requests the item or the item perishes. We assume that perishing is exponentially distributed with mean $1/\theta_k$. The machine chooses whether to produce an item or whether to idle. In case the machine chooses to produce an item, a decision on which class-k item to produce has to be made. This problem, without the possibility of items to perish, was considered in Veatch *et al.* [91]. We denote by $N_k(t)$ the number of class-k items stocked in the inventory. The Markov decision problem is characterized by the following transition rates:

$$b_k^a(m_k) = \mu_k a \text{ and } d_k^a(m_k) = \lambda_k + \theta_k m_k, \tag{4.3.5}$$

where $a = 1(a = 0)$ stands for producing (not producing) an item of that class. If a demand for a class-k item arrives when $N_k(t) = 0$, *i.e.*, no class-k items are left in the stock, then the sale is lost. The latter incurs a per lost cost $D_k > 0$. Every time a class-k item perishes, a per item cost, δ_k, is paid. The system incurs a cost $c_k(m, a)$ per unit of time m class-k items are stocked, $a = 0, 1$. Hence, the cost per unit of time incurred per class k is $C_k(N_k(t), a) = c_k(N_k(t), a) + \delta_k\theta_k N_k(t) + \lambda_k D_k \mathbb{1}_{\{N_k(t)=0\}}$. The objective of the relaxed stochastic model is to minimize the average cost criterion, *i.e.*,

$$\mathbb{E}(\tilde{C}_k(N_k^\phi, S_k^\phi(N_k^\phi))) + D_k\lambda_k\pi_k^\phi(0),$$

with $\tilde{C}_k(m, a) = c_k(m, a) + \delta_k\theta_k m$, and N_k^ϕ the steady-state number of class-k customers under policy ϕ.

In this setting an optimal policy of problem (2.3.3) is of threshold type with 1-0 structure. The proof follows from Proposition 2.1.

The steady-state-probabilities $\pi_k^{n_k}(m_k)$ under threshold n_k are given as follows,

$$\begin{aligned}
&\pi_k^{n_k}(m_k) = \frac{\mu_k^{m_k}}{\prod_{i=1}^{m_k}(\lambda_k + \theta_k i)}\pi^{n_k}(0), \text{ for all } m_k \le n_k + 1,\\
&\pi_k^{n_k}(m_k) = 0, \text{ for all } m_k \ge n_k + 2,
\end{aligned}$$

where $\pi_k^{n_k}(0) = \left(\sum_{m=0}^{n_k+1} \frac{\mu_k^m}{\prod_{i=1}^{m}(\lambda_k+\theta_k i)}\right)^{-1}$. We now check that $\pi_k^n(n+1)$ is strictly decreasing in n as required for indexability in Proposition 2.2. We therefore need to prove that $\pi_k^n(n+1) \le \pi_k^{n-1}(n)$ for all $n \ge 0$, that is,

$$\begin{aligned}
&\frac{\mu_k^{n+1}}{\prod_{i=1}^{n+1}(\lambda_k + \theta_k i)}\pi_k^n(0) \le \frac{\mu_k^n}{\prod_{i=1}^{n}(\lambda_k + \theta_k i)}\pi_k^{n-1}(0)\\
\Leftrightarrow\ &\frac{\mu_k}{\lambda_k + \theta_k(n+1)}\sum_{m=0}^{n}\frac{\mu_k^m}{\prod_{i=1}^{m}(\lambda_k + \theta_k i)} \le \sum_{m=0}^{n+1}\frac{\mu_k^m}{\prod_{i=1}^{m}(\lambda_k + \theta_k i)}\\
\Leftrightarrow\ &\sum_{m=1}^{n}\frac{\mu_k^m}{\prod_{i=1}^{m-1}(\lambda_k + \theta_k i)}\left(\frac{1}{\lambda_k + \theta_k m} - \frac{1}{\lambda_k + \theta_k(n+1)}\right) + 1 \ge 0.
\end{aligned}$$

The latter inequality is satisfied due to $(\frac{1}{\lambda_k+\theta_k m} - \frac{1}{\lambda_k+\theta_k(n+1)}) \ge 0$ for all $m \le n$. Hence, from Proposition 2.2 we have that Whittle's index is given by (2.3.6) in case (2.3.6) is non-increasing. Equation (2.3.6) can be numerically computed and verified to be non-increasing. However, it does not help to provide insights into the properties of Whittle's index policy. This is the main motivation to derive the fluid index, which has a tractable closed-form expression.

The fluid dynamics are given by $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \mu_k s_k(t) - \lambda_k - \theta_k m_k$ for all $m_k \ge 0$. We will assume $\lambda_k < \mu_k$ and then by the convention assumed in Section 4.1.2, *i.e.,* $m_k^a = 0$ in case $f_k^a(m_k) < 0$ for all $m_k \ge 0$ and $a = 0, 1$, we have that $m_k^0 = \max\{0, -\lambda_k/\theta_k\} = 0$ and $m_k^1 = \max\{\frac{\mu_k-\lambda_k}{\theta_k}, 0\} = \frac{\mu_k-\lambda_k}{\theta_k}$. In the stochastic system once the server decides to idle no higher states are visited and therefore, under threshold policy n_k, the average cost corresponding to lost sales equals $D_k \lambda_k \pi_k^{n_k}(0)$. This is a cost that is paid on average, and therefore, the bias optimality criterion of the fluid model cannot capture it, since the fluid system under threshold policies never reaches 0 unless that is the equilibrium point. Instead, we incorporate the sale lost cost per unit of time the passive action is chosen, that is, $C_k(m_k, a) := c_k(m_k, a) + \theta_k \delta_k m_k + \lambda_k D_k(1-a)$.

The fluid index is derived in the following proposition.

Proposition 4.11. *Assume $C_k(m_k, a)$ is convex differentiable and non-decreasing, and $C_k(m_k, 1) - C_k(m_k, 0)$ is convex non-decreasing. Then, the fluid index is non-increasing and given by:*

$$w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + \begin{cases} w_k^{(2)}(m_k), & \text{if } m_k \leq (\mu_k - \lambda_k)/\theta_k \\ w_k^{(1)}(m_k), & \text{if } m_k > (\mu_k - \lambda_k)/\theta_k, \end{cases} \tag{4.3.6}$$

where

$$w_k^{(2)}(m_k) = \left(-m_k - \frac{\lambda_k}{\theta_k}\right)\left(\frac{\mathrm{d}C_k(m_k, 1)}{\mathrm{d}m_k} - \frac{\mathrm{d}C_k(m_k, 0)}{\mathrm{d}m_k}\right) - \frac{\mu_k}{\theta_k}\frac{\mathrm{d}C_k(m_k, 0)}{\mathrm{d}m_k},$$
$$w_k^{(1)}(m_k) = -\frac{\mu_k}{\theta_k}\left(\frac{C_k(m_k, 1) - C_k((\mu_k - \lambda_k)/\theta_k, 1)}{m_k - (\mu_k - \lambda_k)/\theta_k}\right).$$

Proof. Equation (4.3.6) follows from Proposition 4.1 and Lemma 4.12. The index being non-decreasing follows from observing that for any convex non-decreasing function $C_k(m_k, 1)$, for $m_k \leq m_k'$, the function $\frac{C_k(m_k', 1) - C_k(m_k, 1)}{m_k' - m_k}$, is non-decreasing in m_k. Also from the fact that $C_k(m_k, 1) - C_k(m_k, 0)$ being convex and non-decreasing implies that $\frac{\mathrm{d}_k C_k(m_k, 1)}{\mathrm{d}m_k} - \frac{\mathrm{d}_k C_k(m_k, 0)}{\mathrm{d}m_k} \geq 0$ and is non-decreasing. □

We now present an optimal solution of problem (4.1.3).

Proposition 4.12. *Assume $C_k(m_k, a)$ is convex differentiable and non-decreasing, and $C_k(m_k, 1) - C_k(m_k, 0)$ is convex non-decreasing. An optimal solution for problem (4.1.3) with transitions rates (4.3.5) is: $s_k(t) = 1$ if $W \leq w_k(m_k)$ and $s_k(t) = 0$ if $W > w_k(m_k)$, where $w_k(m_k)$ is as given in Proposition 4.11.*

Proof. The proof follows by verifying Assumption 4.1. We have $f_k^a(m_k) = \mu_k a - \lambda_k - \theta_k m_k$, for $a \in \{0, 1\}$. Differentiability of $f_k^a(m_k)$ follows directly as well as monotonicity of $\bar{s}_k(\bar{m}_k)$ in $[m_k^0, m_k^1]$. The function $f_k^a(m_k)$ satisfies

$$\frac{\mathrm{d}^2 f_k^a(m_k)}{\mathrm{d}m_k^2} = 0,$$

for all $m_k \geq 0$. Hence, $f_k^a(m_k)$ is convex in m_k for $a = 0, 1$.

We have $1 - \bar{s}_k(\bar{m}_k) = 1 - (\lambda_k + \theta_k m_k)/\mu_k$, hence

$$\frac{\mathrm{d}^2}{\mathrm{d}m_k^2}\left(1 - \bar{s}_k(\bar{m}_k)\right) = 0,$$

for all $\bar{m}_k \in [m_k^0, m_k^1]$. Hence, $1 - \bar{s}_k(\bar{m}_k)$ is convex in $\bar{m}_k$.

The inequality $f_k^1(m_k) - f_k^1(\bar{m}_k) \leq (\geq) f_k^0(m_k) - f_k^0(\bar{m}_k)$ simplifies to $\bar{m}_k - m_k \leq (\geq) \bar{m}_k - m_k$, which is satisfied for all m_k and $\bar{m}_k$.

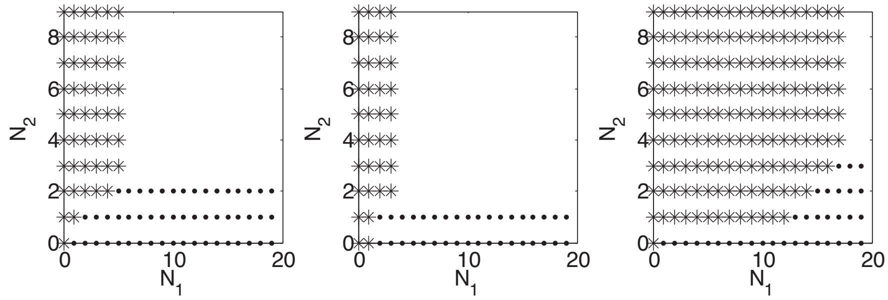


Figure 4.4: In the area with "*" (".") class 1 (class 2) is prioritized and in the white area users are blocked. Left: Optimal policy. Middle: Whittle index policy and, right: the fluid index policy.

Table 4.3: Example 1: Relative sub optimality gap in %.

ρ	0.1	1	1.5	2
Whittle index policy	1.204×10^{-4}	0.0036	0.0099	0.0285
Fluid index policy	0.0532	0.9766	2.8506	4.6276
ρ	2.5	3	3.5	4
Whittle index policy	0.0676	0.1321	0.2242	4.0901
Fluid index policy	2.5092	0.0814	0.0987	4.1111

Then from Proposition 4.1 and Proposition 4.9, an optimal solution for problem (4.1.3) is $s_k(t) = 1$ if $W \leq w_k(m_k)$ and $s_k(t) = 0$ if $W > w_k(m_k)$. □

The fluid index being non-increasing implies that the more class-k items are in stock, less is the priority to produce a class-k item. If the fluid index is negative for all the classes then the fluid index policy prescribes not to produce any item, and hence the machine idles.

We evaluate the performance of the index policies in the following example.

Example 3. In this example we assume two classes of items which are produced at rate $\mu_1 = 4$ and $\mu_2 = 5$, in case the machine decides to produce. The cost function is such that $C_k(m_k, a) = C_k(m_k) + \delta_k\theta_k m_k + \lambda_k D_k \pi^{m_k}(0)$, and we assume $C_k(m_k) = c_k m_k^2 + b_k m_k$ where $c_1 = 1, c_2 = 2$ and $b_1 = b_2 = 1$, and $\theta_1 = 2, \theta_2 = 2.5$. We further assume that the cost for perishing items is $\delta_1 = 0.5, \delta_2 = 3$ and cost per lost sale $D_1 = 20$, $D_2 = 14$. Demands for items arrive at rates $\lambda_1 = 3.5, \lambda_2 = 4.8$. We set $M = 1$, that is, only one machine can produce the items. In Figure 4.4 we illustrate the optimal strategy together with the Whittle index and the fluid index policies for $\rho = 2.5$. The Whittle index policy and the fluid index policy capture the qualitative structure of the optimal solution. We see in Table 4.3 that for several workloads, the relative suboptimality gap of both heuristics is very small. This suggests that any policy that makes the right decisions when either one of the classes is close to empty or empty would behave near optimally.

4.4 Appendix

4.4.1 Proof of Proposition 4.1

We drop the dependency on k throughout the proof. Without loss of generality we assume $m^0 > m^1$. For the sake of clarity we further assume $f^1(m^1) = 0$. The proof for the case $f^1(m^1) < 0$ follows similarly, however, in that case only Cases 2 and 3 (as defined below) have to be analyzed (see below for Cases).

We first introduce the steps followed to prove Proposition 4.1:

- **Step 1:** We prove $w(m)$ to be continuous in m.
- **Step 2:** We prove convexity of $EC(\bar{s}, W)$ with respect to $\bar{s}$. This allows us to characterize the optimal equilibrium point (m^*, s^*), for a given W.
- **Step 3:** We prove that the policy corresponding to $w(m)$ satisfies the Hamilton-Jacobi-Bellman equation, Bertsekas [22, Proposition 3.2.1], using the characterization of W (proven in Step 2) and that, $w(m)$ is continuous (proven in Step 1).

Step 1: Let us first prove that $w(m)$ is a continuous function. To do so let us prove that $\lim_{m\uparrow m^1} w^{(1)}(m) = \lim_{m\downarrow m^1} w^{(2)}(m)$. We observe that

$$\begin{aligned}\lim_{m\uparrow m^1} w^{(1)}(m) =&(f^0(m^1) - f^1(m^1)) \lim_{m\uparrow m^1-} \frac{\big(C(m^1,1) - C(m,1)\big)}{f^1(m)} = \left(\frac{\mathrm{d}f^1(m)}{\mathrm{d}m}\bigg|_{m=m^1} \right)^{-1} \\ &\cdot (f^1(m^1) - f^0(m^1)) \left(\frac{\mathrm{d}C(m,1)}{\mathrm{d}m}\bigg|_{m=m^1} \right),\end{aligned}$$

where in the last equality we applied l'Hopitals rule. Observing that $f^1(m^1) = 0$ we obtain

$$\lim_{m\uparrow m^1} w^{(1)}(m) = -f^0(m^1) \left(\frac{\mathrm{d}C(m,1)}{\mathrm{d}m}\bigg|_{m=m^1} \right) \cdot \left(\frac{\mathrm{d}f^1(m)}{\mathrm{d}m}\bigg|_{m=m^1} \right)^{-1}.$$

We also have

$$\lim_{m\downarrow m^1} w^{(2)}(m) = -f^0(m^1) \left(\frac{\mathrm{d}C(m^1,1)}{\mathrm{d}m^1} \right) \cdot \left(\frac{\mathrm{d}f^1(m^1)}{\mathrm{d}m^1} \right)^{-1},$$

so both limits coincide. Following the same type of arguments it can be proven that $\lim_{m\uparrow m^0} w^{(2)}(m) = \lim_{m\downarrow m^0} w^{(0)}(m)$, which concludes the proof of continuity.

Step 2: We characterize the optimal equilibrium point $(\bar{m}, \bar{s})$, for a given W. We recall that $m^0 > m^1$ and $f^1(m^1) = 0$. Hence, an equilibrium point $(\bar{m}, \bar{s})$ of $\frac{\mathrm{d}m(t)}{\mathrm{d}t}$ is such that

$$(1 - \bar{s})f^0(\bar{m}) + \bar{s}f^1(\bar{m}) = 0, \tag{4.4.1}$$

with $\bar{s} \in [0, 1]$, that is, in equilibrium a fraction $\bar{s}$ of time the action active is chosen and a fraction $1 - \bar{s}$ of time the action passive is chosen. Recall the definition

$$EC(\bar{s}, W) := (1 - \bar{s})C(\bar{m}, 0) + \bar{s}C(\bar{m}, 1) - (1 - \bar{s})W.$$

We will prove this function to be convex in $\bar{s}$, that is, $\frac{\mathrm{d}^2 EC(\bar{s},W)}{\mathrm{d}\bar{s}^2} \geq 0$. After some algebra we observe that

$$\begin{aligned}\frac{\mathrm{d}^2 EC(\bar{s},W)}{\mathrm{d}\bar{s}^2} &= \left(\frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}} - \frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}}\right) 2\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}} + (1-\bar{s})\left(\frac{\mathrm{d}^2 C(\bar{m},0)}{\mathrm{d}\bar{m}^2}\left(\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\right)^2 + \frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}}\frac{\mathrm{d}^2\bar{m}}{\mathrm{d}\bar{s}^2}\right) \\ &\quad + \bar{s}\left(\frac{\mathrm{d}^2 C(\bar{m},1)}{\mathrm{d}\bar{m}^2}\left(\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\right)^2 + \frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}}\frac{\mathrm{d}^2\bar{m}}{\mathrm{d}\bar{s}^2}\right). \end{aligned} \tag{4.4.2}$$

From (4.4.1), we obtain $\bar{s} = f^0(\bar{m})/(f^0(\bar{m}) - f^1(\bar{m}))$ and hence

$$\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}} = \left(\frac{\mathrm{d}\bar{s}}{\mathrm{d}\bar{m}}\right)^{-1} = \left(\frac{f^0(\bar{m})\frac{\mathrm{d}f^1(\bar{m})}{\mathrm{d}\bar{m}} - f^1(\bar{m})\frac{\mathrm{d}f^0(\bar{m})}{\mathrm{d}\bar{m}}}{(f^0(\bar{m}) - f^1(\bar{m}))^2}\right)^{-1} \leq 0, \tag{4.4.3}$$

for all $\bar{s} \in [0,1]$. The inequality follows since both f^0 and f^1 are non-increasing functions, and $f^1(\bar{m}) \leq 0$ and $f^0(\bar{m}) \geq 0$ for all $\bar{m} \in [m^1, m^0]$. From $\mathrm{d}\bar{m}/\mathrm{d}\bar{s} \leq 0$ and $\mathrm{d}^2\bar{m}/\mathrm{d}\bar{s}^2 = -\mathrm{d}^2\bar{s}/\mathrm{d}\bar{m}^2\,(\mathrm{d}\bar{m}/\mathrm{d}\bar{s})^3$, which can be obtained using the Chain Rule, we have that $\mathrm{d}^2\bar{s}/\mathrm{d}\bar{m}^2 \geq 0 \Leftrightarrow \mathrm{d}^2\bar{m}/\mathrm{d}\bar{s}^2 \geq 0$. By assumption, $\bar{s} = f^0(\bar{m})/(f^0(\bar{m}) - f^1(\bar{m}))$ is convex in $\bar{m}$, hence $\mathrm{d}^2\bar{m}/\mathrm{d}\bar{s}^2 \geq 0$. Having $\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}} \leq 0$, $\frac{\mathrm{d}^2\bar{m}^2}{\mathrm{d}\bar{s}^2} \geq 0$, and together with the assumptions (1) $C(m,a)$, $a = 0,1$, is a convex non-decreasing function in m, and (2) $\frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}} \leq \frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}}$, we obtain that (4.4.2) is larger than or equal to 0 and hence $EC(\bar{s},W)$ is convex in $\bar{s} \in [0,1]$.

Due to the function $EC(\bar{s},W)$ being convex in $\bar{s} \in [0,1]$, we have that the optimal equilibrium point will fall into one of the following three options:

- Case 1: $\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \leq 0$ for all $\bar{s} \in [0,1]$, hence the optimal equilibrium point satisfies $s^* = 1, m^* = m^1$.
- Case 2: $\frac{\mathrm{d}EC(s^*,W)}{\mathrm{d}s^*} = 0$, hence the optimal equilibrium point satisfies $s^* \in [0,1]$, $m^* \in [m^1, m^0]$.
- Case 3: $\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \geq 0$ for all $\bar{s} \in [0,1]$, hence the optimal equilibrium point satisfies $s^* = 0$, $m^* = m^0$.

Now note that $\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \geq 0$, if and only if

$$W \geq C(\bar{m},0) + C(\bar{m},1) + (1-\bar{s})\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}} + \bar{s}\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}}, \tag{4.4.4}$$

which after substitution of $(1-\bar{s})f^0(\bar{m}) + \bar{s}f^1(\bar{m}) = 0$ and the expression for $\mathrm{d}\bar{m}/\mathrm{d}\bar{s}$ given in (4.4.3), gives that (4.4.4) is equivalent to

$$W \geq C(\bar{m},0) - C(\bar{m},1) + \frac{(f^1(\bar{m}) - f^0(\bar{m}))(f^0(\bar{m})\frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}} - f^1(\bar{m})\frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}})}{f^0(\bar{m})\frac{\mathrm{d}f^1(\bar{m})}{\mathrm{d}\bar{m}} - f^1(\bar{m})\frac{\mathrm{d}f^0(\bar{m})}{\mathrm{d}\bar{m}}},$$

that is,

$$W \geq C(\bar{m},0) - C(\bar{m},1) + w^{(2)}(\bar{m}).$$

Hence, in Case 3 the W is such that $W \geq C(\bar{m},0) - C(\bar{m},1) + w^{(2)}(\bar{m})$ for all $\bar{m} \in [m^1, m^0]$, and in particular $W \geq w(m^0)$.

Similarly, being in Case 1 implies $W \leq w(m^1)$.

In Case 2, from $\mathrm{d}EC(s^*,W)/\mathrm{d}s^* = 0$ we obtain, $W = C(m^*,0) - C(m^*,1) + w^{(2)}(m^*) = w(m^*)$, for $m^* \in [m^1, m^0]$.

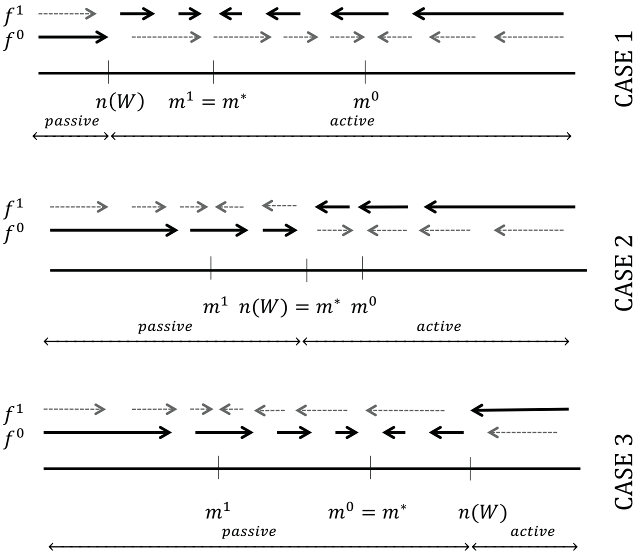


Figure 4.5: The policy $n(W)$: Case 1 above, Case 2 in the middle and Case 3 below. Thick arrows indicating the actual drifts chosen (for each of the cases).

Step 3: Since $s(t) \in \{0, 1\}$ for all t, a sufficient condition for a trajectory to be optimal is to satisfy the HJB Equation, Bertsekas [22, Proposition 3.2.1], it can also be found in Appendix A.3. Rewriting (4.1.4) we get the following condition for all m:

$$0 = \min\{\mathcal{J}_0(m, W), \mathcal{J}_1(m, W)\}, \tag{4.4.5}$$

where

$$\mathcal{J}_0(m, W) = C(m, 0) - W - EC^*(W) + f^0(m)\frac{\partial J(m, W)}{\partial m}, \tag{4.4.6}$$

$$\mathcal{J}_1(m, W) = C(m, 1) - EC^*(W) + f^1(m)\frac{\partial J(m, W)}{\partial m}, \tag{4.4.7}$$

and the function $J(m, W)$ is continuous and differentiable.

For a given W, we consider the policy that prescribes to be passive, $s(t) = 0$, in all states m for which $W \geq w(m)$, and active, $s(t) = 1$, in all states m for which $W < w(m)$. Observe that due to $w(m)$ being non-decreasing, this will be a threshold policy. That is, there exists $n(W) \in \mathbb{Z}_+$ for which $W > w(m)$ for all $m \leq n(W)$ and $W \leq w(m)$ if $m \geq n(W)$, see Figure 4.5. We refer to this policy as threshold policy $n(W)$. We want to prove that the policy $n(W)$ satisfies the HJB (4.4.5). To do so let us define $J^{n(W)}(m, W)$ for a given W as the cost under policy $n(W)$, starting at state m and up to equilibrium, that is,

$$\begin{aligned} J^{n(W)}(m, W) = &\int_0^{t_0} \Big(C(m^{n(W)}(t), s_0) - W(1 - s_0) - EC^*(W)\Big)\mathrm{d}t \\ &+ \int_{t_0}^{\infty} \Big(C(m^{n(W)}(t), s_1) - W(1 - s_1) - EC^*(W)\Big)\,\mathrm{d}t, \end{aligned} \tag{4.4.8}$$

where $s_0 = s(0)$, $s_1 = 1 - s_0$, and $t_0 \geq 0$, the time at which threshold $n(W)$ is reached. Note that $s_0 = 0$ if $m(0) = m \leq n(W)$ and $s_0 = 1$ otherwise. The function $J^{n(W)}(m,W)$ can be written as the sum of two terms, the first term corresponding to the phase from the starting point m up to the time the threshold is reached, t_0. In this phase the control equals s_0. Once the threshold is reached, a switch in the control happens and therefore the second term corresponds to the phase from the switch time t_0 until the equilibrium in reached. In this phase the control equals s_1. This is due to threshold policies having at most one switch in the control, see Figure 4.5.

We now compute $\partial J^{n(W)}(m,W)/\partial m$. We obtain that if $m \leq n(W)$

$$\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W) - C(m,0) + W}{f^0(m)}, \tag{4.4.9}$$

and if $m > n(W)$

$$\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W) - C(m,1)}{f^1(m)}. \tag{4.4.10}$$

This can be seen as follows. Assume $m \leq n(W)$, which implies $s_0 = 0$ and $s_1 = 1$, then from (4.4.8)

$$\begin{aligned}\frac{\partial J^{n(W)}(m,W)}{\partial m} =& \frac{\mathrm{d}t_0}{\mathrm{d}m}\Big(C(n(W),0) - W - EC^*(W) - C(n(W),1) + EC^*(W)\Big) \\ &+ \int_0^{t_0} \frac{\mathrm{d}C(m^{n(W)}(t),0)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}m}\mathrm{d}t + \int_{t_0}^{\infty} \frac{\mathrm{d}C(m^{n(W)}(t),1)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}m}\mathrm{d}t.\end{aligned} \tag{4.4.11}$$

Policy $n(W)$ implies $\mathrm{d}m^{n(W)}(t)/\mathrm{d}t = f^0(m^{n(W)}(t))$ for all $t \in [0, t_0]$. Hence,

$$\int_m^{n(W)} \frac{\mathrm{d}m^{n(W)}(t)}{f^0(m^{n(W)}(t))} = t_0,$$

which implies

$$\begin{aligned}\frac{\mathrm{d}t_0}{\mathrm{d}m} &= \frac{\mathrm{d}}{\mathrm{d}m}\int_m^{n(W)} \frac{\mathrm{d}m^{n(W)}(t)}{f^0(m^{n(W)}(t))} = -\frac{1}{f^0(m)} + \int_m^{n(W)} \frac{\mathrm{d}}{\mathrm{d}m}\left(\frac{1}{f^0(m^{n(W)}(t))}\right)\mathrm{d}m^{n(W)}(t) \\ &= -\frac{1}{f^0(m)},\end{aligned} \tag{4.4.12}$$

since $\mathrm{d}(1/f^0(m^{n(W)}(t)))/\mathrm{d}m = 0$. From $\mathrm{d}m^{n(W)}(t)/\mathrm{d}t = f^0(m^{n(W)}(t))$ we have $\mathrm{d}m/\mathrm{d}t = f^0(m)$. Substituting this together with (4.4.12) in Equation (4.4.11), we obtain (4.4.9). Equation (4.4.10) follows similarly.

For all $m \leq n(W)$, the action under policy $n(W)$ is to keep the bandit passive. In addition, when substituting $\frac{\partial J^{n(W)}(m,W)}{\partial m}$ in (4.4.6), we obtain $\mathcal{J}_0(m,W) = 0$. In order for the threshold policy $n(W)$ to satisfy the HJB in (4.4.5), we therefore need to prove that $\mathcal{J}_1(m,W) \geq 0$. Substituting $\frac{\partial J^{n(W)}(m,W)}{\partial m}$ in (4.4.7) we obtain that this is equivalent to

$$\mathcal{J}_1(m,W) \geq 0 \Leftrightarrow W \geq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}(C(m,1) - EC^*(W)), \tag{4.4.13}$$

for all $m \notin [m^1, m^0]$ with $m \leq n(W)$, and

$$\mathcal{J}_1(m,W) \geq 0 \Leftrightarrow W \leq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}(C(m,1) - EC^*(W)), \tag{4.4.14}$$

for all $m \in [m^1, m^0]$ with $m \leq n(W)$. If (4.4.13) is satisfied for $m \notin [m^1, m^0]$ and (4.4.14) for $m \in [m^1, m^0]$, then the action under policy $n(W)$, to keep the bandit passive, is optimal.

Assume now $m > n(W)$. Hence, action under policy $n(W)$ is to keep the bandit active. Substituting $\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W) - C(m,1)}{f^1(m)}$ in (4.4.7) we then obtain $\mathcal{J}_1(m,W) = 0$. In order for the threshold policy $n(W)$ to satisfy the HJB in (4.4.5), we need therefore to prove that $\mathcal{J}_0(m,W) \geq 0$. Substituting $\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W) - C(m,1)}{f^1(m)}$ in (4.4.6), this is equivalent to

$$\mathcal{J}_0(m,W) \geq 0 \Leftrightarrow W \leq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}(C(m,1) - EC^*(W)), \tag{4.4.15}$$

for all $m > n(W)$. If (4.4.15) is satisfied for all $m > n(W)$ then the action under policy $n(W)$, to keep the bandit active, is optimal.

Hence, if conditions (4.4.13)$-$(4.4.15) are satisfied, then threshold policy $n(W)$ is optimal. It remains to be proved that conditions (4.4.13)$-$(4.4.15) are satisfied. This will be done in the remainder of the proof for the three different cases.

Let us first assume that $m^* = m^1$ and $W \leq w(m^1)$, that is, Case 1 (as proven in Step 2). Hence $EC^*(W) = C(m^1, 1)$. Recall that threshold policy $n(W)$ implies that $W \geq w(m)$ for all $m \leq n(W)$ and $W \leq w(m)$ if $m \geq n(W)$. Hence, $W \leq w(m^1)$ and $w(m)$ being non-decreasing imply that $n(W) \leq m^1$, see Figure 4.5. Conditions (4.4.13)$-$(4.4.15) reduce then to the following: the HJB is satisfied if and only if $W \geq (\leq) C(m,0) - C(m,1) + w^{(1)}(m)$ for all $m \leq (\geq) n(W)$. This is equivalent to $W \geq (\leq) w(m)$ for all $m \leq (\geq) n(W)$, since $w^{(1)}(m)$ is non-decreasing and $W \leq w(m^1)$. Hence, in Case 1 the threshold policy $n(W)$ satisfies the HJB and is hence optimal.

Similarly, if $m^* = m^0$ and $W \geq w(m^0)$, that is, Case 3, then $EC^*(W) = C(m^0, 0) - W$. Since under threshold policy $n(W)$, $W \geq w(m)$ for all $m \leq n(W)$ and $W \leq w(m)$ if $m \geq n(W)$, $w(m)$ being non-decreasing implies $n(W) \geq m^0$, see Figure 4.5. Using $EC^*(W) = C(m^0, 0) - W$, we obtain that conditions (4.4.13)$-$(4.4.15) simplify to $W \geq (\leq) C(m,0) - C(m,1) + w^{(0)}(m)$, for all $m \leq (\geq) n(W)$. This is equivalent to $W \geq (\leq) w(m)$ for all $m \leq (\geq) n(W)$, due to $w^{(0)}(m)$ being non-decreasing and $W \geq w(m^0)$. Hence, in Case 3, threshold policy $n(W)$ satisfies the HJB and is hence optimal.

We are left with Case 2 in which W is such that $\frac{\mathrm{d}E(s^*,W)}{\mathrm{d}s^*} = 0$, and $s^* \in [0,1]$, that is, $W = w(m^*)$. Hence $n(W) = m^*$, see Figure 4.5, by definition of $n(W)$. In this setting we have that

$$EC^*(W) = (1 - s^*)(C(m^*, 0) - W) + s^* C(m^*, 1).$$

Substituting the latter in Conditions (4.4.13) and (4.4.15) the conditions simplify to

$$W \geq (\leq) C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}\Big(C(m,1) - (1 - s^*)(C(m^*,0) - W) - s^* C(m^*, 1)\Big), \tag{4.4.16}$$

for all $m \leq m^1$($m \geq m^0$).

Condition (4.4.14) and (4.4.15) reduce to

$$W \leq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}\Big(C(m,1) - (1-s^*)(C(m^*,0) - W) - s^*C(m^*,1)\Big), \quad (4.4.17)$$

for all $m \in [m^1, m^*]$ and

$$W \leq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}\Big(C(m,1) - (1-s^*)(C(m^*,0) - W) - s^*C(m^*,1)\Big), \quad (4.4.18)$$

for all $m \in [m^*, m^0]$.

Taking into account that $f^1(m) \geq 0$, for all $m < m^1$, and $f^1(m) \leq 0$, otherwise, and that by assumption $f^0(m)(1-s^*) + s^*f^1(m) > 0$, for all $m < m^*$, and $f^0(m)(1-s^*) + s^*f^1(m) < 0$ for all $m^* > m$, Conditions (4.4.16)–(4.4.18) reduce to the following:

$$\begin{aligned} W \geq (\leq)\Big(C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}\Big(C(m,1) - (1-s^*)C(m^*,0) - s^*C(m^*,1)\Big)\Big) \\ \cdot \frac{f^1(m)}{f^0(m)(1-s^*) + s^*f^1(m)}, \end{aligned} \quad (4.4.19)$$

for all $m < m^*(m > m^*)$. Since $W = w(m^*)$, and $w^{(2)}(\cdot)$ and $w(\cdot)$ are non-decreasing, in order to prove (4.4.19) it suffices to verify that:

- RHS in (4.4.19) $\xrightarrow{m \to m^*} C(m^*,0) - C(m^*,1) + w^{(2)}(m^*)$,
- RHS in (4.4.19) $\leq C(m^*,0) - C(m^*,1) + w^{(2)}(m^*)$ for all $m < m^*$,
- RHS in (4.4.19) $\geq C(m^*,0) - C(m^*,1) + w^{(2)}(m^*)$ for all $m > m^*$.

We have proven these conditions in Lemma 4.2 (Section 4.4.2) and this therefore concludes the proof of Proposition 4.1.

4.4.2 Lemma 4.2

Lemma 4.2. *Assume $m_k^0 > m_k^1$. Let $(m_k^*, s_k^*) \in [m_k^1, m_k^0] \times [0,1]$, be the optimal equilibrium point of* (4.1.3). *Define*

$$\begin{aligned} F_k(m_k) := &\frac{f_k^1(m_k)}{f_k^0(m_k)(1-s_k^*) + s_k^* f_k^1(m_k)}\Big(C_k(m_k,0) - C_k(m_k,1) \\ &+ \frac{f_k^1(m_k) - f_k^0(m_k)}{f_k^1(m_k)}\Big(C_k(m_k,1) - (1-s_k^*)C_k(m_k^*,0) - s_k^*C_k(m_k^*,1)\Big)\Big), \end{aligned}$$

for all m_k. Then,

$$\lim_{m_k \to m_k^*} F_k(m_k) = C_k(m_k^*,0) - C_k(m_k^*,1) + w_k^{(2)}(m_k^*) = w_k(m_k^*), \quad (4.4.20)$$

$$F_k(m_k) \leq C_k(m_k^*,0) - C_k(m_k^*,1) + w_k^{(2)}(m_k^*), \text{ for all } m_k < m_k^*, \quad (4.4.21)$$

$$F_k(m_k) \geq C_k(m_k^*,0) - C_k(m_k^*,1) + w_k^{(2)}(m_k^*) \text{ for all } m_k > m_k^*. \quad (4.4.22)$$

Proof. In Section 4.4.2 we prove (4.4.20), and in Section 4.4.2, we prove (4.4.21) and (4.4.22).

We drop the dependency on k throughout the proof.

Proof of (4.4.20)

After some algebra we observe that $F(m)$ reduces to

$$\frac{f^1(m)C(m,0)-f^0(m)C(m,1)}{f^1(m)s^*+f^0(m)(1-s^*)}-\frac{(f^1(m)-f^0(m))(1-s^*)C(m^*,0)}{f^1(m)s^*+f^0(m)(1-s^*)}-\frac{(f^1(m)-f^0(m))s^*C(m^*,1)}{f^1(m)s^*+f^0(m)(1-s^*)},$$

which further simplifies to

$$C(m^*,0)-C(m^*,1)+\frac{f^1(m)(C(m,0)-C(m^*,0))}{f^1(m)s^*+f^0(m)(1-s^*)}+\frac{f^0(m)(C(m^*,1)-C(m,1))}{f^1(m)s^*+f^0(m)(1-s^*)}. \tag{4.4.23}$$

We have

$$\begin{aligned}
&\lim_{m\to m^*}(f^1(m)-f^0(m))\frac{(f^1(m)s^*+f^0(m)(1-s^*))}{m^*-m}=\left(\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}s^*+\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}(1-s^*)\right)(f^1(m^*)-f^0(m^*))\\
&=-\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}s^*f^1(m^*)+\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}s^*f^0(m^*)-\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}(1-s^*)f^1(m^*)+\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}(1-s^*)f^0(m^*),
\end{aligned}$$

where in the first step we used that $s^*f^1(m^*)=(s^*-1)f^0(m^*)$ and applied L'Hopitals rule. Substituting $s^*f^1(m^*)=(s^*-1)f^0(m^*)$ in the first and fourth terms we obtain

$$\lim_{m\to m^*}(f^1(m)-f^0(m))\frac{(f^1(m)s^*+f^0(m)(1-s^*))}{m^*-m}=\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}f^0(m^*)-\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}f^1(m^*).$$

If we now substitute this last obtained expression in (4.4.23), where we multiply and divide the fraction terms by $(f^1(m)-f^0(m))/(m^*-m)$, we obtain that as $m\to m^*$ Equation (4.4.23) reduces to

$$\begin{aligned}
&C(m^*,0)-C(m^*,1)+\frac{(f^1(m^*)-f^0(m^*))(f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}-f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*})}{f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}-\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}f^1(m^*)},\\
&=C(m^*,0)-C(m^*,1)+w^{(2)}(m^*),
\end{aligned}$$

since $\lim_{m\to m^*}\frac{C(m,a)-C(m^*,a)}{m^*-m}=-\frac{\mathrm{d}C(m^*,a)}{\mathrm{d}m^*}$, for $a=0,1$. This concludes the proof of (4.4.20).

Proof of (4.4.21) **and** (4.4.22)

We want to prove that $F(m)\leq(\geq)C(m^*,0)-C(m^*,1)+w^{(2)}(m^*)$ for all $m<m^*(m>m^*)$. After substitution of $w^{(2)}(m^*)$ together with (4.4.23) this reduces to

$$\begin{aligned}
&\frac{f^1(m)(C(m,0)-C(m^*,0))}{f^1(m)s^*+f^0(m)(1-s^*)}+\frac{f^0(m)(C(m^*,1)-C(m,1))}{f^1(m)s^*+f^0(m)(1-s^*)}\\
&\leq(\geq)\frac{(f^1(m^*)-f^0(m^*))(f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}-f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*})}{f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}-\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}f^1(m^*)},
\end{aligned} \tag{4.4.24}$$

for all $m<(>)m^*$. In the latter we substitute $s^*=f^0(m^*)/(f^0(m^*)-f^1(m^*))\geq0$, and we divide both sides by $f^0(m^*)-f^1(m^*)$, which we recall is positive due to $f^1(m^*)\leq0$ and $f^0(m^*)\geq0$. Also, under the assumption that $f^a(\cdot)$ is non-increasing for $a=0,1$ and $\bar{s}(\bar{m})$ strictly monotone in $\bar{m}$ we have that

$\nexists \bar{m} \in [m^1, m^0]$ such that $f^0(\bar{m}) = f^1(\bar{m}) = 0$. Then, Inequality (4.4.24) writes

$$\frac{f^1(m)\,(C(m,0) - C(m^*,0))}{f^0(m^*)(f^1(m) - f^1(m^*)) - f^1(m^*)(f^0(m) - f^0(m^*))} + \frac{f^0(m)\,(C(m^*,1) - C(m,1))}{f^0(m^*)(f^1(m) - f^1(m^*)) - f^1(m^*)(f^0(m) - f^0(m^*))}$$
$$\leq (\geq) \frac{f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*} - f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}}{f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*} - f^1(m^*)\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}},$$

for all $m < (>) m^*$. After multiplying and dividing by $m - m^*$ in the left hand side, it simplifies to

$$\frac{-f^1(m)\left(\frac{C(m,0)-C(m^*,0)}{m-m^*}\right)}{f^1(m^*)\frac{f^0(m)-f^0(m^*)}{m-m^*} - f^0(m^*)\frac{f^1(m)-f^1(m^*)}{m-m^*}} + \frac{f^0(m)\left(\frac{C(m,1)-C(m^*,1)}{m-m^*}\right)}{f^1(m^*)\frac{f^0(m)-f^0(m^*)}{m-m^*} - f^0(m^*)\frac{f^1(m)-f^1(m^*)}{m-m^*}}$$
$$\leq (\geq) \frac{f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*}}{f^1(m^*)\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*} - f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}}, \tag{4.4.25}$$

for all $m < (>) m^*$. The function $f^a(\cdot)$ is convex non-increasing for $a = 0, 1$. Hence,

$$\frac{f^1(m) - f^1(m^*)}{m - m^*} \leq (\geq) \frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}, \text{ and}$$
$$\frac{f^0(m) - f^0(m^*)}{m - m^*} \leq (\geq) \frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*},$$

for all $m < (>) m^*$. Since $f^1(m^*) \leq 0$ and $f^0(m^*) \geq 0$, it follows that

$$f^1(m^*)\frac{f^0(m) - f^0(m^*)}{m - m^*} - f^0(m^*)\frac{f^1(m) - f^1(m^*)}{m - m^*} \geq (\leq) f^1(m^*)\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*} - f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*},$$

for all $m < (>) m^*$, that is, the denominator of (4.4.25) is ordered. Also note that both denominators are strictly positive. We now will show that the corresponding ordering holds as well for the nominators of (4.4.25), that is,

$$-f^1(m)\left(\frac{C(m,0) - C(m^*,0)}{m - m^*}\right) + f^0(m)\left(\frac{C(m,1) - C(m^*,1)}{m - m^*}\right)$$
$$\leq (\geq) f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*}, \tag{4.4.26}$$

for all $m < m^* (m > m^*)$. This will conclude the proof.

Let us first consider $m \in [m^1, m^*) (m \in (m^*, m^0])$. Then

$$-f^1(m)\left(\frac{C(m,0) - C(m^*,0)}{m - m^*}\right) + f^0(m)\left(\frac{C(m,1) - C(m^*,1)}{m - m^*}\right)$$
$$\leq (\geq) - f^1(m)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*} + f^0(m)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*} \leq (\geq) - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*} + f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}$$

for all $m \in [m^1, m^*] (m \in [m^*, m^0])$. The first inequality holds due to $C(\cdot, a)$ being a convex non-decreasing function and therefore $\frac{C(m,a)-C(m^*,a)}{m-m^*} \leq (\geq) \frac{\mathrm{d}C(m^*,a)}{\mathrm{d}m^*}$ for $a = 0, 1$ and $m < (>) m^*$, and $-f^1(m), f^0(m) \geq 0$ for all m. The second inequality follows from the assumptions $\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*} \geq \frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}$, and $f^1(m) - f^1(m^*) \geq (\leq) f^0(m) - f^0(m^*)$, for all $m < (>) m^*$.

Let us now consider $m \leq m^1$. Then

$$
\begin{aligned}
&- f^1(m)\left(\frac{C(m,0)-C(m^*,0)}{m-m^*}\right) + f^0(m)\left(\frac{C(m^*,1)-C(m,1)}{m-m^*}\right)\\
&\leq (f^0(m)-f^1(m))\left(\frac{C(m,1)-C(m^*,1)}{m-m^*}\right) \leq (f^0(m^*)-f^1(m^*))\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m}\\
&\leq f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*},
\end{aligned}
$$

for all $m \leq m^1$, where the first inequality follows from the fact that $\left(\frac{C(m,0)-C(m^*,0)}{m-m^*}\right) \geq \left(\frac{C(m,1)-C(m^*,1)}{m-m^*}\right)$ for all $m \leq m^1 \leq m^*$ and $-f^1(m) \leq 0$ for all $m \leq m^1$. The second inequality follows from the assumption $f^0(m)-f^0(m^*) \leq f^1(m)-f^1(m^*)$ for all $m \leq m^1$ and $\frac{C(m,1)-C(m^*,1)}{m-m^*} \leq \frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}$ for all $m \leq m^*$. Finally, the third inequality follows from $\mathrm{d}C(m^*,0)/\mathrm{d}m^* \geq \mathrm{d}C(m^*,1)/\mathrm{d}m^*$ and $-f^1(m^*) \geq 0$.

Similarly we have

$$
- f^1(m)\left(\frac{C(m,0)-C(m^*,0)}{m-m^*}\right) + f^0(m)\left(\frac{C(m,1)-C(m^*,1)}{m-m^*}\right) \geq f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*},
$$

for all $m \geq m^0$, using that $f^0(m) < 0$, for all $m \geq m^0$ and $\left(\frac{C(m,0)-C(m^*,0)}{m-m^*}\right) \geq \left(\frac{C(m,1)-C(m^*,1)}{m-m^*}\right)$ for all $m \geq m^0 \geq m^*$. Hence (4.4.26) holds for all m, which concludes the proof.

□

4.4.3 Proof of Proposition 4.2

Let us drop the dependency on k throughout the proof. We compute the index as $\lambda \downarrow 0$ for 0-1 type of threshold policies first and for 1-0 type of threshold policies later on.

0-1 type of threshold policies. Assume that the optimal threshold policy that solves (2.3.3) is of 0-1 type. Recall the assumption $b^a(m) = \lambda\gamma$ for $a = 0, 1$. In order to lighten the notation throughout the proof we define

$$
\overline{D}^n(m) := \begin{cases} d^0(1)\cdot d^0(2)\cdot \ldots \cdot d^0(m), & \forall m \leq n,\\ d^0(1)\cdot \ldots \cdot d^0(n)\cdot d^1(n+1)\cdot \ldots \cdot d^1(m), & \forall m > n. \end{cases} \tag{4.4.27}
$$

Observe that $\overline{D}^n(m) = \overline{D}^{n-1}(m)$ for all $m \leq n-1$. This observation will be used throughout the proof.

We consider birth-and-death processes and therefore $\pi^n(m) := \lambda^m\gamma^m\pi^n(0)/\overline{D}^n(m)$, where $\pi^n(0) = \left(\sum_{m=0}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}\right)^{-1}$. We now compute the expression of $\pi^{n-1}(0)/\pi^n(0)$, which will be used later in the proof. We have

$$
\frac{\pi^{n-1}(0)}{\pi^n(0)} = \frac{\sum_{m=0}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}{\sum_{m=0}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}} = 1 + \frac{\sum_{m=n}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)} - \sum_{m=n}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}}{\sum_{m=0}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}}, \tag{4.4.28}
$$

the latter as $\lambda \to 0$ reduces to $1 + \lim_{\lambda\downarrow 0}\frac{\mathcal{O}(\lambda^n)}{1+\mathcal{O}(\lambda)} = 1$.

We have that $W(n)$ as given in Equation (2.3.6) can be written as

$$W(n) = \frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))}, \tag{4.4.29}$$

where

$$\begin{aligned}
\xi_1(n) &:= \sum_{m=0}^{n-1} C(m,0)(\pi^n(m)-\pi^{n-1}(m)),\\
\xi_2(n) &:= C(n,0)\pi^n(n) - C(n,1)\pi^{n-1}(n),\\
\xi_3(n) &:= \sum_{m=n+1}^{\infty} C(m,1)(\pi^n(m)-\pi^{n-1}(m)).
\end{aligned} \tag{4.4.30}$$

Let us first analyze the first term. That is,

$$\begin{aligned}
\frac{\xi_1(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} &= \frac{\sum_{m=0}^{n-1} C(m,0)\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}(\pi^n(0)-\pi^{n-1}(0))}{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)}\pi^n(0)+(\pi^n(0)-\pi^{n-1}(0))\sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}\\
&= \frac{\sum_{m=0}^{n-1} C(m,0)\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)}\frac{1}{1-\frac{\pi^{n-1}(0)}{\pi^n(0)}}+\sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}.
\end{aligned} \tag{4.4.31}$$

We now substitute the expression obtained in (4.4.28) in (4.4.31). Then, as $\lambda \downarrow 0$, the denominator in (4.4.31) becomes

$$\begin{aligned}
\lim_{\lambda\downarrow 0} -\frac{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)}\sum_{m=0}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}}{\sum_{m=n}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}-\sum_{m=n}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}}+\sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)} &= \lim_{\lambda\downarrow 0} -\frac{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)}(1+\mathcal{O}(\lambda))}{\left(\frac{\gamma^n}{\overline{D}^n(n)}-1\right)(\lambda^n+\mathcal{O}(\lambda^{n+1}))}+1+\mathcal{O}(\lambda)\\
&= -\left(1-\frac{d^0(n)}{d^1(n)}\right)^{-1}+1 = \frac{d^0(n)}{d^0(n)-d^1(n)}.
\end{aligned}$$

Substituting the latter in (4.4.31) we obtain that as $\lambda \downarrow 0$

$$\frac{\xi_1(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} = C(0,0)\left(1-\frac{d^1(n)}{d^0(n)}\right)+\mathcal{O}(\lambda). \tag{4.4.32}$$

Let us now analyze the second term, that is,

$$\frac{\xi_2(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} = \frac{C(n,0)-C(n,1)\frac{d^0(n)}{d^1(n)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}{1+(1-\frac{\pi^{n-1}(0)}{\pi^n(0)})\frac{\overline{D}^n(n)}{\lambda^n\gamma^n}\sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}. \tag{4.4.33}$$

We substitute (4.4.28) in the latter equation. As $\lambda \downarrow 0$ the denominator in (4.4.33) simplifies to

$$\lim_{\lambda\downarrow 0} 1 - \frac{\frac{\overline{D}^n(n)}{\lambda^n\gamma^n}\left(\sum_{m=n}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)} - \sum_{m=n}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}\right)}{\sum_{m=0}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}} \sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)} = \lim_{\lambda\downarrow 0} 1 - \frac{1-\frac{d^0(n)}{d^1(n)}+\mathcal{O}(\lambda)}{1+\mathcal{O}(\lambda)}(1+\mathcal{O}(\lambda))$$
$$= \frac{d^0(n)}{d^1(n)},$$

and substituting the latter in (4.4.33) we obtain that as $\lambda \downarrow 0$

$$\begin{aligned}\frac{\xi_2(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} &= C(n,0)\frac{d^1(n)}{d^0(n)} - C(n,1)\\ &= C(n,0)-C(n,1)+C(n,0)\left(\frac{d^1(n)}{d^0(n)}-1\right).\end{aligned} \tag{4.4.34}$$

To conclude the proof we obtain the expression of the third term, that is,

$$\frac{\xi_3(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} = \frac{\sum_{m=n+1}^{\infty} C(m,1)\left(\frac{\lambda^m\gamma^m}{\overline{D}^n(m)} - \frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}\frac{\pi^{n-1}(0)}{\pi^n(0)}\right)}{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)} + (1-\frac{\pi^{n-1}(0)}{\pi^n(0)})\sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}.$$

The latter, letting $\lambda \downarrow 0$, reduces to $\frac{\mathcal{O}(\lambda^{n+1})}{\mathcal{O}(\lambda)+1}$, and hence

$$\lim_{\lambda\downarrow 0}\frac{\xi_3(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))} = 0.$$

Summing the latter together with Equations (4.4.32) and (4.4.34) we obtain $W(n) = W^{LT}(n) + o(1)$, as $\lambda \downarrow 0$, where

$$W^{LT}(n) = C(n,0) - C(n,1) + (C(n,0)-C(0,0))\left(\frac{d^1(n)}{d^0(n)}-1\right).$$

This concludes the proof.

1-0 type of threshold policies. Assume that the optimal threshold policy that solves (2.3.3) is of 1-0 type now. Recall the assumption $b^a(m) = \lambda\gamma a$ for $a = 0, 1$. In order to lighten the notation throughout the proof we define

$$\overline{D}^n(m) := \begin{cases} d^1(1)\cdot d^1(2)\cdot \ldots \cdot d^1(m), & \forall m \le n,\\ d^1(1)\cdot \ldots \cdot d^1(n)\cdot d^0(n+1), & m = n+1,\end{cases} \tag{4.4.35}$$

when the threshold is considered to be n. Observe that $\overline{D}^n(m) = \overline{D}^{n-1}(m)$ for all $m \le n-1$. This observation will be used throughout the proof.

In the framework of this thesis we consider birth-and-death processes and therefore

$$\pi^n(m) := \lambda^m\gamma^m\pi^n(0)/\overline{D}^n(m),$$

for all $m \le n+1$, and 0 otherwise, where $\pi^n(0) = \left(\sum_{m=0}^{n+1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}\right)^{-1}$.

We have that $W(n)$ as given in Equation (2.3.6) reduces to

$$W(n) = \frac{\xi_1(n) + \xi_2(n) + \xi_3(n)}{\pi^n(n+1) - \pi^{n-1}(n)}, \tag{4.4.36}$$

where

$$\begin{aligned}
\xi_1(n) &:= \sum_{m=0}^{n-1} C(m,1)(\pi^n(m) - \pi^{n-1}(m)), \\
\xi_2(n) &:= C(n,1)\pi^n(n) - C(n,0)\pi^{n-1}(n), \\
\xi_3(n) &:= C(n+1,0)\pi^n(n+1),
\end{aligned} \tag{4.4.37}$$

and $\pi^n(m)$ the steady-state probability of being in state m.

We now compute the expression of $\pi^{n-1}(0)/\pi^n(0)$, which will be used later in the proof. We hence have

$$\begin{aligned}
\frac{\pi^{n-1}(0)}{\pi^n(0)} &= \frac{\sum_{m=0}^{n+1} \frac{\lambda^m \gamma^m}{\overline{D}^n(m)}}{\sum_{m=0}^{n} \frac{\lambda^m \gamma^m}{\overline{D}^{n-1}(m)}} = 1 + \frac{\frac{\lambda^n \gamma^n}{\overline{D}^n(n)} + \frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^{n+1}(n+1)} - \frac{\lambda^n \gamma^n}{\overline{D}^{n-1}(n)}}{\sum_{m=0}^{n} \frac{\lambda^m \gamma^m}{\overline{D}^{n-1}(m)}} \\
&= 1 + \frac{\frac{\lambda^n \gamma^n}{\overline{D}^n(n-1)} \left(\frac{1}{d^1(n)} - \frac{1}{d^0(n)} \right) + \frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^{n+1}(n+1)}}{\sum_{m=0}^{n} \frac{\lambda^m \gamma^m}{\overline{D}^{n-1}(m)}}.
\end{aligned} \tag{4.4.38}$$

The latter as $\lambda \to 0$ reduces to $1 + \lim_{\lambda \downarrow 0} \frac{\mathcal{O}(\lambda^n)}{1+\mathcal{O}(\lambda)} = 1$.

Let us first analyze the first term. That is,

$$\frac{\xi_1(n)}{\pi^n(n+1) - \pi^{n-1}(n)} = \frac{\sum_{m=0}^{n-1} C(m,1) \frac{\lambda^m \gamma^m}{\overline{D}^n(m)} \left(1 - \frac{\pi^{n-1}(0)}{\pi^n(0)} \right)}{\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)} - \frac{\lambda^n \gamma^n}{\overline{D}^{n-1}(n)} \frac{\pi^{n-1}(0)}{\pi^n(0)}} \tag{4.4.39}$$

We now substitute the expression obtained in (4.4.38) in (4.4.39). Then, since

$$\begin{aligned}
&\frac{1 - \frac{\pi^{n-1}(0)}{\pi^n(0)}}{\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)} - \frac{\lambda^n \gamma^n}{\overline{D}^{n-1}(n)} \frac{\pi^{n-1}(0)}{\pi^n(0)}} \\
&= - \frac{\frac{\lambda\gamma}{d^0(n+1)d^1(n)} + \frac{1}{d^1(n)} - \frac{1}{d^0(n)}}{\sum_{m=0}^{n} \frac{\lambda^m \gamma^m}{\overline{D}^{n-1}(m)}} \cdot \frac{1}{\frac{\lambda\gamma}{d^0(n+1)d^1(n)} - \frac{1}{d^0(n)} - \frac{\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)} + \frac{\lambda^n\gamma^n}{\overline{D}^{n-1}(n-1)}\left(\frac{1}{d^1(n)} - \frac{1}{d^0(n)}\right)}{\sum_{m=0}^{n} \frac{\lambda^m \gamma^m}{\overline{D}^{n-1}(m)}}} \\
&= \frac{d^0(n)}{d^1(n)} - 1 + o(\lambda),
\end{aligned}$$

where the last equality is obtained by letting $\lambda \to 0$. Substituting the later in (4.4.39) we obtain that as $\lambda \downarrow 0$

$$\frac{\xi_1(n)}{\pi^n(n+1) - \pi^{n-1}(n)} = C(0,1) \left(\frac{d^0(n)}{d^1(n)} - 1 \right) + \mathcal{O}(\lambda). \tag{4.4.40}$$

Let us now analyze the second term, that is,

$$\frac{\xi_2(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = \frac{C(n,1)\frac{d^0(n)}{d^1(n)} - C(n,0)\frac{\pi^{n-1}(0)}{\pi^n(0)}}{\frac{\lambda\gamma d^0(n)}{d^0(n+1)d^1(n)} - \frac{\pi^{n-1}(0)}{\pi^n(0)}}. \tag{4.4.41}$$

We substitute (4.4.38) in the latter equation. As $\lambda \downarrow 0$, (4.4.41) simplifies to

$$\frac{\xi_2(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = -C(n,1)\frac{d^0(n)}{d^1(n)} + C(n,0) + \mathcal{O}(\lambda).$$

Hence,

$$\lim_{\lambda\downarrow 0}\frac{\xi_2(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = C(n,0) - C(n,1) + C(n,1)\left(1-\frac{d^0(n)}{d^1(n)}\right). \tag{4.4.42}$$

To conclude the proof we obtain the expression of the third term, that is,

$$\frac{\xi_3(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = \frac{C(n+1,0)\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)}}{\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)} - \frac{\lambda^n}{\overline{D}^{n-1}(n)}\frac{\pi^{n-1}(0)}{\pi^n(0)}} = \frac{C(n+1,0)}{1-\frac{\overline{D}^n(n+1)}{\lambda\gamma\overline{D}^{n-1}(n)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}.$$

The latter, letting $\lambda \downarrow 0$, reduces to $\frac{\mathcal{O}(\lambda)}{\mathcal{O}(\lambda)+1}$, and hence

$$\lim_{\lambda\downarrow 0}\frac{\xi_3(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = 0.$$

Summing the latter together with Equations (4.4.40) and (4.4.42) we obtain $W(n) = W^{LT}(n) + o(1)$, as $\lambda \downarrow 0$, where

$$W^{LT}(n) = C(n,0) - C(n,1) + (C(n,1) - C(0,1))\left(1-\frac{d^0(n)}{d^1(n)}\right).$$

Which concludes the proof.

4.4.4 Proof of Proposition 4.3

We drop the dependency on k throughout the proof.

We have $d^a(m) > 0$ for $a = 0,1$ and for all $m > 0$ and $b^a(m) = \lambda\gamma$ for all m and $a = 0,1$ when 0-1 type of threshold policy solves (2.3.3) optimally. Similarly, $d^a(m) > 0$ for $a = 0,1$ and all $m > 0$ and $b^a(m) = \lambda\gamma a$ for all m_k and $a = 0,1$ when 1-0 type of threshold policy solves (2.3.3) optimally.

As $\lambda \downarrow 0$ $f^a(m) \to -d^a(m)$, and by assumption $d^a(m) > 0$ for all m, then if m is such that $-d^a(m) = 0$ this implies $m < 0$. Hence, by the convention adopted in Section 4.1.2 this implies $m^a = 0$ for $a = 0,1$. The latter together with Proposition 4.1 gives us that the fluid index policy for the 0-1 type of threshold policies is given by $w(m) = C(m,0) - C(m,1) + w^{(0)}(m)$ for all m and for the 1-0 type of threshold policies by $w(m) = C(m,0) - C(m,1) + w^{(1)}(m)$ for all m with $m^0 = m^1 = 0$. Then, from Proposition 4.2 we have

$$\lim_{\lambda\downarrow 0} W(m) = C(m,0) - C(m,1) + \frac{C(m,a) - C(0,a)}{d^a(m)}\left(d^1(m) - d^0(m)\right) = \lim_{\lambda\downarrow 0} w(m).$$

This concludes the proof.

4.4.5 Proof of Proposition 4.4

We drop the dependency on k throughout the proof.

As $n \to \infty$, then the fluid index is given by $w(n) = C(n,0) - C(n,1) + \delta(\mu+\theta') - \delta'\theta' + w^{(3)}(n)$. We have assumed that $C(n,a)$, $a = 0,1$, are upper bounded by a polynomial of degree P. Therefore, we can write $C(n,a) = E(n,a) + o(1)$, for large values of n, where $E(n,1) = \sum_{i=0}^{P} C^{(P,i)} n^i$, with

$$C^{(P,i)} := \lim_{n\to\infty} \frac{C(n,1) - \sum_{j=i+1}^{P} C^{(P,j)} n^j}{n^i},$$

and $E(n,0) = \sum_{i=0}^{Q} E^{(Q,i)} n^i$, with

$$E^{(Q,i)} := \lim_{n\to\infty} \frac{C(n,0) - \sum_{j=i+1}^{Q} E^{(Q,j)} n^j}{n^i},$$

Then, as $n \to \infty$, $w(n) = w^\infty(n) + o(1)$, where $w^\infty(n) = \delta(\mu+\theta') - \delta'\theta' + w^c(n) + o(1)$, and

$$w^c(n) = E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta}\frac{(E(n,0) - E(\lambda/\theta,0))}{n - \lambda/\theta}.$$

Note that $(E(n,0) - E(\lambda/\theta,0))/(n-\lambda/\theta)$ for large values of n can equivalently be written as

$$\begin{aligned}
&\frac{\sum_{i=0}^{Q} E^{(Q,i)} n^i - \sum_{i=0}^{Q} E^{(Q,i)} (\lambda/\theta)^i}{n - \lambda/\theta} = \sum_{i=0}^{Q} E^{(Q,i)} \frac{(n^i - (\lambda/\theta)^i)}{n-\lambda/\theta} = \sum_{i=2}^{Q} E^{(Q,i)} \left(\sum_{j=0}^{i} \left(\frac{\lambda}{\theta}\right)^j n^{i-1-j} \right) \\
&= \frac{E(n,0)}{n} + \frac{E^{(Q,1)}\left(\frac{\lambda}{\theta}\right) + E^{(Q,2)}\left(\frac{\lambda}{\theta}\right)^2 + \ldots + E^{(Q,Q)}\left(\frac{\lambda}{\theta}\right)^Q}{n} + \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j}\left(\frac{\lambda}{\theta}\right)^{j+1} \\
&= \frac{E(n,0)}{n} + \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j}\left(\frac{\lambda}{\theta}\right)^{j+1} + o(1). \qquad (4.4.43)
\end{aligned}$$

We then compute $\lim_{n\to\infty} W(n)/w(n)$, which by the result in (3.3.4) is equivalent to

$$\begin{aligned}
&\lim_{n\to\infty}\frac{W(n)}{w(n)} = \lim_{n\to\infty}\frac{W^\infty(n) + o(1)}{w^\infty(n) + o(1)} = \lim_{n\to\infty}\frac{\delta(\mu+\theta') - \delta'\theta' + W^c(n) + o(1)}{\delta(\mu+\theta') - \delta'\theta' + w^c(n) + o(1)} \\
&= \lim_{n\to\infty}\frac{E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta}\left(\frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j}\left(\frac{\lambda}{\theta}\right)^{j+1}\right) + \mathcal{O}(1)}{E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta}\left(\frac{E(n,0)}{n} + \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j}\left(\frac{\lambda}{\theta}\right)^{j+1}\right) + \mathcal{O}(1)} \\
&= 1 + o(1),
\end{aligned}$$

which follows from the fact that both in the denominator and numerator the highest term comes from $E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta}\frac{E(n,0)}{n}$. This concludes the proof for the expression in (4.2.1).

Let us now obtain the expression in (4.2.2) with the extra assumptions $P = Q$ and $C^{(P,i)} = E^{(P,i)}$ for all $i \in \{2, \ldots, P\}$. Observe that under this assumption we obtain from (4.4.43) that $(E(n,0)-E(\lambda/\theta,0))/(n-\lambda/\theta)$, for large values of n, can be written as

$$\frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} + o(1),$$

and hence,

$$\begin{aligned} w^{\infty}(n) =& \delta(\mu+\theta') - \delta'\theta' + E(n,0) - E(n,1) \\ &+ \frac{(\mu+\theta'-\theta)}{\theta}\left(\frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j}\left(\frac{\lambda}{\theta}\right)^{j+1}\right) + o(1). \end{aligned}$$

Then, by the result in (3.3.4) we have $W^{\infty}(n) = w^{\infty}(n) + o(1)$, and hence $W(n) = w(n) + o(1)$ for large values of n which concludes the proof for (4.2.2).

Chapter 5

Optimal fluid control for an abandonment queue

Contents

5.1 Introduction **105**
5.2 Model description **107**
5.3 Fluid control model **108**
5.4 Heuristics for an arbitrary number of classes **114**
5.5 Numerical results **116**
5.6 Appendix **120**

In this chapter we investigate an abandonment queue with linear holding cost. This is a particular case of Chapter 3, where convex holding cost were considered. We focus on two different settings: (1) customers can abandon both while waiting in the queue and while being served, (2) only customers that are in the queue can abandon. In Chapter 3 we used the Lagrangian relaxation approach to derive a heuristic. In this chapter, we propose a fluid control model as a deterministic approximation. For an overload scenario (*i.e.*, $\rho > 1$) we obtain that the $\tilde{c}\mu/\theta$ rule optimally solves the fluid control problem. For an underload scenario (*i.e.*, $\rho < 1$) we use Pontryagin's Minimum Principle to obtain the optimal control of the fluid model for two classes of customers: there exists a switching curve that splits the two-dimensional state-space into two regions such that when the number of customers in both classes is sufficiently small the optimal policy follows the $\tilde{c}\mu$-rule and when the number of customers is sufficiently large the optimal policy follows the $\tilde{c}\mu/\theta$-rule. The same structure is observed in the optimal policy of the stochastic model for an arbitrary number of classes. Based on this we develop a heuristic and by numerical experiments we evaluate its performance and compare it to several index policies. We observe that the suboptimality gap of our solution is small.

5.1 Introduction

In this chapter we investigate the fundamental question of how to share *one* common resource among multiple classes of customers in the presence of abandonments. As it has been discussed in the introduction, and in Chapter 3 determining the exact optimal policy is very challenging. In Chapter 3 we have performed the Lagrangian relaxation approach to obtain a heuristic for the original model. In this chapter we will adopt the fluid scaling method. We can find several works in the literature regarding efficient control in abandonment queues under limiting regimes. One approach is to study the system in the Halfin-Whitt heavy-traffic regime. That is, the total arrival rate and the number of servers both become large in such a way that the traffic intensity approaches one. For the abandonment queue this was first studied in Garnett *et al.* [45]. This scaling gives rise to a diffusion control problem, for which the optimal controls are investigated in Atar *et al.* [10], Harrison *et al.* [55] and shown to be state dependent. In an overload setting the abandonment queue has been studied under a fluid scaling in Atar *et al.* [8, 9], where the authors scale the number of servers and the arrival rate and show that the $\tilde{c}\mu/\theta$ rule (*i.e.*, the policy where strict priority is given according to the indices $\tilde{c}\mu/\theta$) is asymptotically fluid optimal (here $\tilde{c}$ is the holding plus abandonment cost, θ is the abandonment rate and μ the service rate). The overload assumption is crucial in their analysis, since under this assumption the trajectories of the fluid model converge to a strictly positive state which completely characterizes the performance under the average performance criteria. The $\tilde{c}\mu/\theta$-rule emerges naturally as the policy that optimizes the performance associated to this absorbing state. Without abandonments, the $\tilde{c}\mu$-rule, i.e., strict priority is given according to the indices $\tilde{c}\mu$, is optimal in a multi-class single server queue for average reward and discounted cost criteria, in the preemptive and non-preemptive cases, see for example Buyukkoc *et al.* [33].

The stochastic model we adopt here will be that of Chapter 3 with linear holding cost. The analysis will concentrate on two different stochastic models: (1) customers can abandon when they are waiting in the queue and also while they are being served, (2) customers that are in the queue can abandon but customers in service cannot. With the notation introduced in Chapter 3 case (1) corresponds to $\theta' = \theta$, and case (2) to $\theta' = 0$. In an overload setting, we determine the optimal equilibrium point and show that under the $\tilde{c}\mu/\theta$ rule the dynamics converge to this point, which in fact is non-zero. In the underload case the fluid model will empty in finite time, hence we will seek the optimal trajectory that minimizes the cost of draining the fluid. The latter makes the analysis considerably harder than in the overload case. Using PMP we characterize the optimal solution for the case with two classes of customers. The optimal solution has a remarkable structure, there exists a switching curve that splits the two-dimensional state-space into two regions such that: when the number of customers is sufficiently small the optimal policy follows the $\tilde{c}\mu$-rule and when the number of customers is sufficiently large the optimal policy follows the $\tilde{c}\mu/\theta$-rule. Recall that in Chapter 3 we obtained that priority had to be given according to $\tilde{c}\mu/\theta$-rule. Solving the optimal stochastic control problem numerically we observe this same behavior (priorities depend on $\tilde{c}\mu$ and $\tilde{c}\mu/\theta$) where the shape of the switching curve is very well approximated by the one found in the fluid model. In fact, the combination of the $\tilde{c}\mu$ rule and the $\tilde{c}\mu/\theta$ rule is also observed numerically in the optimal stochastic control for more than two classes. We use this insight to propose a heuristic for the stochastic model (for an arbitrary number of classes). At last, by numerical experiments we evaluate the performance of the fluid-based heuristic and several index policies, among which the Whittle's index policy derived in Chapter 3, and observe that the suboptimality gap of the solution proposed in this chapter is rather small. We emphasize here that the heuristic proposed in this chapter works well across all loads, while the index policies $\tilde{c}\mu$ and $\tilde{c}\mu/\theta$, although being rather easy to implement, achieve only good performance in either the underload or the overload setting.

The remainder of the chapter is organized as follows. In Section 5.2 we present the stochastic model with abandonments and the optimization problem. In Section 5.3 we introduce the related fluid model and solve its fluid control problem for the underload case (for two classes of customers) and for the overload case. In Section 5.4 we develop a heuristic for the stochastic model for an arbitrary number of classes and in Section 5.5 we numerically compare the performance of the fluid-based heuristic with that of the optimal policy and several index policies that have been derived and discussed in Chapter 3. Most of the proofs can be found in Appendix 5.6.

5.2 Model description

We consider a multi-class single-server queue with K classes of customers. Class-k customers arrive according to a Poisson process with rate λ_k and have an exponentially distributed service requirement with mean $1/\mu_k$. A class-k customer can abandon the system after an exponentially distributed amount of time with mean $1/\theta_k$ if the customer is waiting in the queue, and with mean $1/\theta'_k$ if the customer is already receiving service. We define $\rho_k = \lambda_k/\mu_k$ as the traffic load of class k and $\rho = \sum_k \rho_k$ as the total load. We assume that the server has capacity 1 and can serve at most one customer at a time, where the service can be preemptive. At every decision epoch, a policy ϕ decides which class is served. Because of the Markov property we can focus on policies that base decisions on the current number of customers present in the various classes. For a given policy ϕ, the control variable $(S_1^\phi(t), \ldots, S_K^\phi(t))$ denotes the class of the customer that is in service at time t, *i.e.*, if at time t class k is in service, then $S_k^\phi(t) = 1$ and $S_l^\phi(t) = 0$ for $l \neq k$. Hence, it satisfies $S_k^\phi(t) \in \{0, 1\}$ and $\sum_{k=1}^K S_k^\phi(t) \leq 1$.

We are interested in two different models, depending on whether or not a customer in service becomes impatient and hence can abandon:

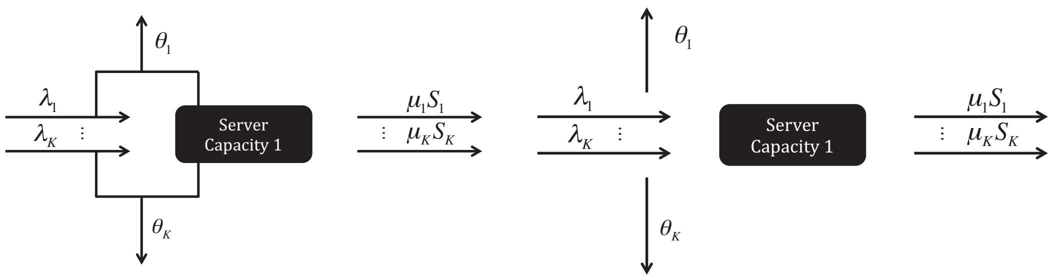
- The $\theta = \theta'$ case: class-k customers can abandon both while waiting in the queue and while being served at rate θ_k, $k \in \{1, \ldots, K\}$, see Figure 5.1a. Hence, $\theta_k = \theta'_k$. We further assume $\delta_k = \delta'_k$ for all $k \in \{1, \ldots, K\}$.
- The $\theta' = 0$ case: customers can abandon only while waiting in the queue, see Figure 5.1b. Hence, $\theta'_k = 0$ for all $k \in \{1, \ldots, K\}$.

Both models have been studied in the literature, *e.g.*, in Down *et al.* [41] the $\theta = \theta'$ case is studied, while the authors of Atar *et al.* [8, 9], Ayesta *et al.* [15] consider the $\theta' = 0$ case. In the analysis of Chapter 3 both settings have been considered.

For a given policy ϕ, let $N_k^\phi(t)$ denote either the number of class-k customers in the system (the $\theta = \theta'$ case) or the number of class-k customers in the queue (the $\theta' = 0$ case). Let c_k denote the holding cost per unit time for class-k customers. Let δ_k denote the cost for each class-k customer that abandons. Our objective is to minimize the average cost, that is,

$$\min_\phi \limsup_{T\to\infty} \sum_{k=1}^K \frac{1}{T}\mathbb{E}\left(\int_0^T c_k N_k^\phi(t)\ \mathrm{d}t + \delta_k R_k^\phi(T)\right) = \min_\phi \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \sum_{k=1}^K \tilde{c}_k N_k^\phi(t)\ \mathrm{d}t\right),$$

where $R_k^\phi(T)$ denotes the number of class-k customers that abandoned in the interval $[0, T]$, $\tilde{c}_k := c_k + \delta_k\theta_k$, $k = 1, \ldots, K$, and we used that $\mathbb{E}(R_k^\phi(T)) = \theta_k \mathbb{E}(\int_0^T N_k^\phi(t)\mathrm{d}t)$. We note that for $\theta' = 0$ case only waiting customers can abandon and can hence contribute to the abandonment cost. In addition, implicitly we



(a) The $\theta = \theta'$ case (b) The $\theta' = 0$ case

Figure 5.1: Multi-class single-server queue with abandonments.

assumed that for the $\theta' = 0$ case, only waiting customers contribute to the holding cost[1]. For the $\theta = \theta'$ case, all customers will have a contribution to the abandonment cost (a customer in service can abandon) and we have assumed that all customers contribute to the holding cost.

The above described stochastic control problems have proved to be very difficult to solve, as we have argued at the end of Section 3.1.

5.3 Fluid control model

In this section the stochastic models ($\theta = \theta'$, $\theta' = 0$) presented in Section 5.2 are approximated by the deterministic fluid model, where only the mean dynamics are taken into account, see Section 1.3.2. That is, let $m_k(t)$ be the amount of class-k fluid and $s_k(t)$ the control parameter. Then the fluid dynamics is described by the following set of differential equations:

$$\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \lambda_k - \mu_k s_k(t) - \theta_k m_k(t), \text{ for all } k \in \{1, \ldots, K\},$$

$$(s_1(t), \ldots, s_K(t)) \in \mathcal{S}, \; m_k(t) \geq 0, \text{ for all } k \in \{1, \ldots, K\}, \text{ for all } t,$$

with

$$\mathcal{S} := \{s = (s_1, \ldots, s_K) \text{ s.t } \sum_{k=1}^{K} s_k \leq 1, \, s_k \geq 0, \text{ for all } k \in \{1, \ldots, K\}\}.$$

For the fluid analysis we will make a distinction between two different scenarios: (1) $\rho < 1$, which we refer to as the underload and (2) $\rho > 1$, which we refer to as the overload setting. Note that in case $\rho < 1$, any non-idling control (*i.e.*, $\sum_{k=1}^{K} s_k(t) = 1$ if $\sum_{k=1}^{K} m_k(t) > 0$) converges to the equilibrium point $(0, \ldots, 0)$[2]. Hence, when $\rho < 1$ we aim at minimizing the total cost until reaching the equilibrium point $(0, \ldots, 0)$. This can be written as

$$\min_{s(t) \in \mathcal{S}} \int_0^\infty \sum_{k=1}^{K} \tilde{c}_k m_k(t) \, \mathrm{d}t.$$

[1]For the $\theta' = 0$ case, the latter was also assumed in Atar *et al.* [8, 9], while Ayesta *et al.* [15] assumed customers in service contribute to the holding cost as well.

[2]Consider $w(t) := \sum_{k=1}^{K} m_k(t)/\mu_k$. Then $\frac{\mathrm{d}w(t)}{\mathrm{d}t} = \rho - \sum_{k=1}^{K} s_k(t) - \sum_{k=1}^{K} \frac{\theta_k}{\mu_k} m_k(t) < \rho - 1 < 0$, hence $w(t)$ converges to zero.

Note that this last cost coincides with the bias-optimality cost criteria, as considered in Chapter 4, since the cost at equilibrium is 0. When $\rho > 1$, an equilibrium point will necessarily be different than $(0, \ldots, 0)$. Hence, for $\rho > 1$ our objective is to minimize the average cost, *i.e*,

$$\min_{s(t)\in\mathcal{S}} \lim_{T\to\infty} \frac{1}{T} \int_0^T \sum_{k=1}^K \tilde{c}_k m_k(t)\ \mathrm{d}t.$$

Throughout the study we refer to this optimal fluid control problem as Problem P.

5.3.1 Optimal policy in underload for two classes of customers

In this section we assume $\rho < 1$ and solve the fluid control model. We focus on the case of two classes of customers, whose solution is already rather involved. However, it gives us intuition on the structure of the optimal policy for an arbitrary number of classes.

We will see that an optimal policy can be of two possible shapes: either a switching curve emerges, *i.e.*, we prioritize one class above the switching curve and the other class below the switching curve, or one of the two classes is prioritized. This gives us four different type of strategies. As we show in the following proposition, the optimal strategy is fully characterized by the ordering of $\tilde{c}_1\mu_1$ and $\tilde{c}_2\mu_2$ and of $\tilde{c}_1\mu_1/\theta_1$ and $\tilde{c}_2\mu_2/\theta_2$. The proof can be found in Appendix 5.6.2.

Proposition 5.1. *Assume $K = 2$ and let $\lambda_k, \mu_k, \theta_k, c_k$ and δ_k be given for $k \in \{1, 2\}$. Assume $\rho < 1$. If $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$, then an optimal solution $s^*(\cdot)$ for Problem P under the total cost criteria is:*

- *If $\tilde{c}_2\mu_2 \leq \tilde{c}_1\mu_1$, then*
  - $s^* = (0, 1)$ *when* $m_2 > h(m_1)$,
  - $s^* = (1, 0)$ *when* $m_2 \leq h(m_1)$ *and* $m_1 > 0$,
  - $s^* = (\rho_1, 1 - \rho_1)$ *when* $m_2 \leq h(0)$ *and* $m_1 = 0$,

  where the switching curve $h(\cdot)$ is given by

  $$h(m_1) := \frac{a_1 m_1 + a_2 + (a_3 m_1 - a_2)\left(\dfrac{\theta_1 m_1 + \mu_1 - \lambda_1}{\mu_1 - \lambda_1}\right)^{\frac{\theta_2}{\theta_1}}}{a_4 m_1} + \frac{\lambda_2}{\theta_2}, \tag{5.3.1}$$

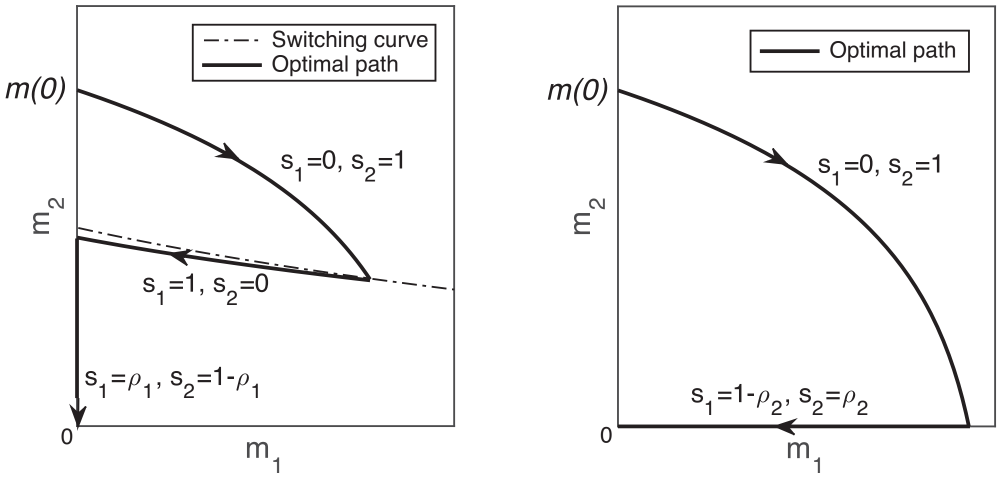
  with

  $$a_1 = \tilde{c}_2 \frac{\mu_2}{\theta_2}(1 - \rho)\,; \quad a_2 = a_1 \frac{\mu_1}{\theta_1}(1 - \rho_1); \quad a_3 = -\left(\tilde{c}_2\frac{\mu_2}{\theta_2} - \tilde{c}_1\frac{\mu_1}{\theta_1}\right)(1 - \rho_1),$$

  $$\text{and } a_4 = \left(\tilde{c}_2\frac{\mu_2}{\theta_2} - \tilde{c}_1\frac{\mu_1}{\theta_1}\right)\frac{\theta_2}{\mu_2}.$$

  That is, serve class 2 until the switching curve $h(\cdot)$ is reached, then serve class 1 until $m_1 = 0$. From that moment on, keep $m_1 = 0$ and give the rest of service to class 2, see Figure 5.2a.

- *If $\tilde{c}_2\mu_2 \geq \tilde{c}_1\mu_1$, then*
  - $s^* = (0, 1)$ *when* $m_2 > 0$,



(a) Optimal path when $\tilde{c}_2\mu_2 \leq \tilde{c}_1\mu_1$.

(b) Optimal path when $\tilde{c}_2\mu_2 \geq \tilde{c}_1\mu_1$.

Figure 5.2: Optimal strategy assuming $\dfrac{\tilde{c}_2\mu_2}{\theta_2} \geq \dfrac{\tilde{c}_1\mu_1}{\theta_1}$, and the optimal path.

– $s^* = (1 - \rho_2, \rho_2)$ *when* $m_2 = 0$.

That is, serve class 2 until $m_2 = 0$*. From that moment on, keep* $m_2 = 0$ *and give the rest of service to class 1, see Figure 5.2b.*

The solution in the case where $\tilde{c}_2\mu_2/\theta_2 \leq \tilde{c}_1\mu_1/\theta_1$ *is equivalent with the indices swapped.*

Remark 5.1 (Arbitrary number of classes)**.** *Given the complexity to find an optimal solution for the fluid control model in underload when* $K = 2$*, we did not aim at obtaining an analytical solution for an arbitrary number of classes* K*. Instead, in Section 5.4 we develop a heuristic using the insights obtained for the case K=2.*

Remark 5.2. *Observe that the switching curve* $h(\cdot)$ *defined in Proposition 5.1 intersects with the vertical axis at*

$$h(0) = (1 - \rho)\frac{\mu_2}{\theta_1\theta_2}\left(\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\dfrac{\tilde{c}_2\mu_2}{\theta_2} - \dfrac{\tilde{c}_1\mu_1}{\theta_1}}\right).$$

Since we are in the underload case, the latter is greater than or equal to 0, if and only if

$$\left(\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2\right) / \left(\frac{\tilde{c}_2\mu_2}{\theta_2} - \frac{\tilde{c}_1\mu_1}{\theta_1}\right) \geq 0.$$

The proof of the above proposition follows from Lemma 5.1. Before presenting this lemma we first provide intuition for the structure of the optimal policy, which is characterized by a very simple rule based on the comparison of the indices $\tilde{c}\mu$ and $\tilde{c}\mu/\theta$. When the amount of fluid is small enough, the $\tilde{c}\mu$ rule is optimal. This can be explained as follows. Note that the derivative of the cost is given by $\sum_{k=1}^{K} \tilde{c}_k \frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \sum_{k=1}^{K} \tilde{c}_k \left(\lambda_k - \mu_k s_k(t) - \theta_k m_k(t)\right)$. The $\tilde{c}\mu$-rule myopically minimizes the derivative and is hence optimal in the short run. Close to the origin this is exactly what the optimal control prescribes.

However, in the long term, one cannot neglect the effect of abandonments. For example, if $\tilde{c}_1\mu_1 > \tilde{c}_2\mu_2$, but $\theta_1 >> \theta_2$, then the myopic rule would prioritize class 1. However, this minimizes $m_1(t)$, which has a negative impact on the derivative of the cost (cf. the term $\theta_1 m_1(t)$). Hence, in the long run it might be good to keep the amount of class-1 fluid high, since class 1 has a high abandonment rate. In Proposition 5.1 we showed that in a state far from the origin, the index that appropriately combines the above described effects is the $\tilde{c}\mu/\theta$ index. We will show in Proposition 5.2 that the $\tilde{c}\mu/\theta$ rule is in fact optimal when $\rho > 1$, *i.e.*, in the overload setting.

The switching curve $h(\cdot)$, as defined in Proposition 5.1, describes the states in which it is optimal to switch from the $\tilde{c}_k\mu_k/\theta_k$ rule to the $\tilde{c}_k\mu_k$ rule. We can learn the following from the formula for $h(\cdot)$:

- As we can see from Remark 5.2, the ratio between $\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2$ and $\dfrac{\tilde{c}_2\mu_2}{\theta_2} - \dfrac{\tilde{c}_1\mu_1}{\theta_1}$ determines $h(0)$, and hence the height of the switching curve. From this we observe that as the difference in the values for the $\tilde{c}\mu$ index grows large (small) relative to that of the $\tilde{c}\mu/\theta$ index, the height of the switching curve grows (goes to zero) and hence the optimal fluid control gets closer to the $\tilde{c}\mu$ rule ($\tilde{c}\mu/\theta$ rule).

- As the traffic load approaches one, *i.e.*, $\rho \uparrow 1$, the switching curve $h(\cdot)$ converges to $\overline{h}(\cdot)$ with $\overline{h}(0) = 0$ and $\overline{h}(m_1) < 0$ for $m_1 > 0$. Hence, the $\tilde{c}\mu/\theta$ rule is optimal for the fluid model as $\rho \uparrow 1$. As we will see in Section 5.3.2, the $\tilde{c}\mu/\theta$ rule is optimal in the overload setting ($\rho > 1$) as well, showing continuity in the optimal solution.

Remark 5.3 (Multi-class queue with deadlines)**.** *In the case $c_k = 0$, $k = 1, \dots, K$, the model becomes a multi-class queue with deadlines: customers need to be served before a deadline that is exponentially distributed with parameter θ_k and in case they do not receive service before their deadline they abandon the queue giving a cost δ_k. In this particular case the $\tilde{c}\mu$ rule reduces to $\delta\mu\theta$ rule and the $\tilde{c}\mu/\theta$ rule reduces to the $\delta\mu$ rule.*

The following lemma is necessary in order to prove Proposition 5.1. Its proof is presented in Appendix 5.6.1. The lemma states that the index $\tilde{c}_k\mu_k$ determines the optimal action when the amount of fluid in both class 1 and class 2 is small.

Lemma 5.1. *Let $K = 2$ and let $m(0) = (\varepsilon, \varepsilon)$ with $\varepsilon > 0$ small enough. If $\rho < 1$ and*

$$\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2 \ (\textit{resp. } \tilde{c}_1\mu_1 \leq \tilde{c}_2\mu_2),$$

then it is optimal to give priority to class 1 (resp. class 2) until the origin is reached.

5.3.2 Optimal policy in overload for an arbitrary number of classes

In this section we assume again an arbitrary number of classes, *i.e.*, $K \geq 2$. To complete the analysis of Problem P we are left with the setting $\rho > 1$, in which case the objective is to minimize the average cost (the latter being strictly positive). The following proposition states an optimal control for the fluid model.

Proposition 5.2. *Let $\lambda_k, \mu_k, \theta_k, c_k$ and δ_k be given for $k \in \{1, \dots, K\}$, and assume the classes are ordered such that $\frac{\tilde{c}_1\mu_1}{\theta_1} \geq \frac{\tilde{c}_2\mu_2}{\theta_2} \geq \dots \geq \frac{\tilde{c}_K\mu_K}{\theta_K}$. If $\rho > 1$, then an optimal solution $s^*(\cdot)$ for Problem P under the average cost criteria is:*

$$s^*(t) = (\rho_1, \dots, \rho_l, 1 - \sum_{i=1}^{l(t)} \rho_i, 0, \dots, 0),$$

with $l(t) := \min\{k : m_{k+1}(t) > 0\}$*. That is, priority is given according to the index* $\tilde{c}\mu/\theta$*.*

Proof. We first determine the optimal equilibrium point. An equilibrium point satisfies $0 = \lambda_k - \mu_k s_k - \theta_k m_k,$ for all k. Hence, the optimal control (in equilibrium) that minimizes the equilibrium point is given by

$$\arg\min_{s\in\mathcal{S}} \sum_{k=1}^{K} \tilde{c}_k m_k = \arg\min_{s\in\mathcal{S}} \sum_{k=1}^{K} \tilde{c}_k \frac{\lambda_k - \mu_k s_k}{\theta_k} = \arg\min_{s\in\mathcal{S}} \frac{\sum_{k=1}^{K} \tilde{c}_k \mu_k}{\theta_k} s_k.$$

This is minimized by giving highest priority according to the $\tilde{c}\mu/\theta$ rule, that is the optimal equilibrium point is given by $m^* = (0, \ldots, 0, \frac{\lambda_{j+1}-\mu_{j+1}(1-\sum_{i=1}^{j}\rho_i)}{\theta_{j+1}}, \frac{\lambda_{j+2}}{\theta_{j+2}}, \ldots, \frac{\lambda_K}{\theta_K})$ and $s^* = (\rho_1, \ldots, \rho_j, 1 - \sum_{i=1}^{j}\rho_i, 0, \ldots, 0)$, with j such that $\sum_{i=1}^{j}\rho_i < 1$ and $\sum_{i=1}^{j+1}\rho_i \geq 1$.

It remains to be checked that under the control $s^*(\cdot)$ as stated in the proposition, the fluid dynamics converges to the optimal equilibrium point. This can be seen as follows. Let $m^*(\cdot)$ denote the trajectory corresponding to the control $s^*(\cdot)$. Consider $w_j^*(t) := \sum_{k=1}^{j} m_k^*(t)/\mu_k$. By definition of $s^*(t)$ we have $\mathrm{d}w_j^*(t)/\mathrm{d}t = \sum_{k=1}^{j}\rho_k - 1 - \sum_{k=1}^{j} m_k^*(t)/\theta_k < -(1 - \sum_{k=1}^{j}\rho_k)$ when $w_j^*(t) > 0$. Hence, in a finite time T the process hits zero, $w_j^*(T) = 0$, and stays there. From that moment on, class $j+1$ is given capacity $1 - \sum_{k=1}^{j}\rho_k$ if present. Hence, it follows directly that this converges to the point m_{j+1}^*, which solves $0 = \lambda_{j+1} - \mu_{j+1}(1 - \sum_{k=1}^{j}\rho_k) - \theta_{j+1}m_{j+1}$. Since for $t > T$ we have $m_{j+1}^*(t) > 0$, classes $j+2, \ldots, K$ do not receive any service. Hence, their dynamics are described by $\mathrm{d}m_i^*(t)/\mathrm{d}t = \lambda_i - \theta_i m_i^*(t)$, and $m_i^*(t)$ converges to λ_i/θ_i, $i \in \{j+2, \ldots, K\}$. □

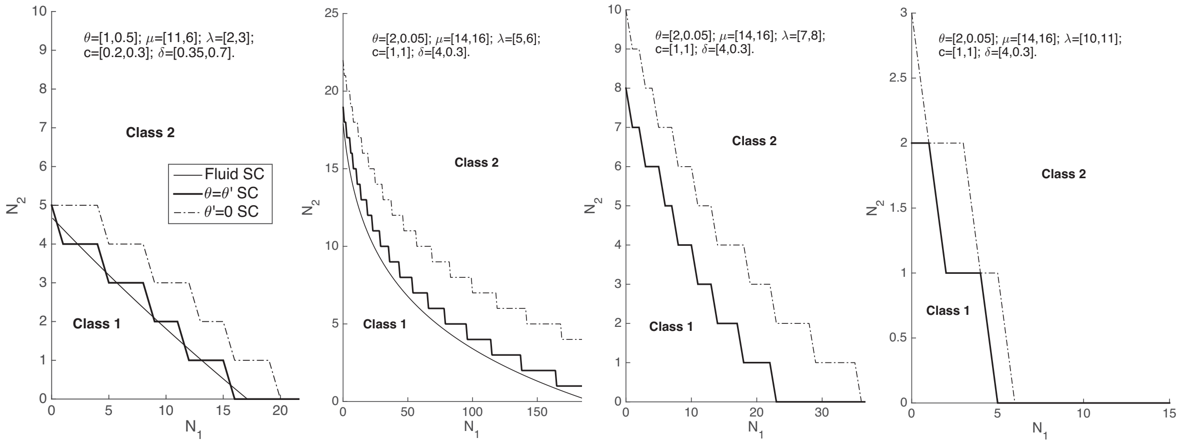
We note that the $\tilde{c}\mu/\theta$ rule has previously been proposed by Atar et al. in [8, 9], where optimal scheduling in the presence of abandonments was studied for the many-server setting. The rule was obtained by solving a fluid control model. The fluid model is similar to the one of Proposition 5.2, but has the additional condition $s_k \leq m_k$, which is due to the multi-server setting. Moreover, $\tilde{c}\mu/\theta$ rule coincides with Whittle's index obtained in Proposition 3.3 for the $c_k' = c_k$, $\theta_k = \theta_k'$ and $\delta_k = \delta_k'$ case.

5.3.3 Optimal control comparison of stochastic model with fluid model

In this section we compare the switching curve that we obtained for the fluid model with the solution for the stochastic models $\theta = \theta'$ and $\theta' = 0$ obtained numerically by value iteration. In Figure 5.3 we make this comparison for different sets of parameters. Note that the optimal stochastic switching curve of the $\theta = \theta'$ case is always below the switching curve of the $\theta' = 0$ case. This is due to the fact that allowing customers to abandon while being served, as in the $\theta = \theta'$ case, makes the effect of abandonments more significant.

We consider the underload case $\rho < 1$ in Figures 5.3a and 5.3b, the critical regime $\rho = 1$ in Figure 5.3c and the overload setting $\rho > 1$ in Figure 5.3d. Moreover, we note that Figures 5.3b, 5.3c and 5.3d correspond to the parameters of Example 1 in Section 5.5.

In Figures 5.3a-5.3b the parameters are such that $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ and $\tilde{c}_1\mu_1/\theta_1 \leq \tilde{c}_2\mu_2/\theta_2$, hence the optimal fluid solution is characterized by a switching curve and priority is given to class 2 above the curve and to class 1 below the curve. We observe that the fluid optimal switching curve approximates the stochastic optimal switching curve very well, except for a constant that apparently disappears after the fluid scaling.



(a) $\rho = 0.6818$. (b) $\rho = 0.7321$. (c) $\rho = 1$. (d) $\rho = 1.4018$.

Figure 5.3: Switching curves for $\theta = \theta'$, $\theta' = 0$ and the fluid control model.

In Figures 5.3c-5.3d the optimal stochastic policy is characterized by a switching curve where class 2 is served in states above the curve and class 1 in states below the curve. The optimal control in the fluid model is however to give strict priority to class 2 (in case $\rho > 1$), since $\tilde{c}_1\mu_1/\theta_1 \le \tilde{c}_2\mu_2/\theta_2$. For Figure 5.3c the average number of customers in the system $\theta = \theta'$ under the optimal policy is $(\overline{m}_1, \overline{m}_2) = (0.7796, 4.1194)$ which lies below the switching curve. Hence, the stochastic optimal policy will give most of the time priority to class 1. The optimal fluid control does not capture this property since the fluid switching curve $h(\cdot)$ vanishes for $\rho = 1$. In Example 1 of Section 5.5 we will see that the suboptimality gap when applying the optimal fluid control to the stochastic model is around 30%. In case $\rho > 1$ is large enough, our index policy turns out to work well, see the numerical Section 5.5. This is explained by the fact that the process is living above the switching curve. For example, for the parameters as chosen in Figure 5.3d, the average number of customers in the system $\theta = \theta'$ is $(\overline{m}_1, \overline{m}_2) = (3.0088, 3.4849)$ which lies above the switching curve. Hence, the stochastic optimal policy will give most of the time priority to class 2, which coincides with the optimal fluid control. In Example 1 of Section 5.5 we will see that the suboptimality gap when applying the optimal fluid control to the stochastic model is small.

In Section 5.4 we discuss how to translate the fluid optimal solution to the stochastic setting. In Section 5.5 we will numerically evaluate the performance of the heuristic when applied to the stochastic model. In fact, we will observe good performance. However, we do not have any result on the suboptimality gap. In the literature, asymptotic fluid optimality results have been obtained for various dynamic scheduling problems in queueing models, see for example Bäuerle [17], Gajrat *et al.* [44], Maglaras [68], Meyn [69], Verloop *et al.* [94]. More precisely, it is shown that when employing the optimal control resulting from the fluid model to the stochastic model, the fluid-scaled cost converges to the optimal cost of the fluid control model, the latter being in fact a provable lower bound on the stochastic cost. In this particular model the fluid model is presented as an approximation, there is no certainty that when applying the optimal fluid control in the stochastic model, this will be asymptotically optimal. We do believe though that when scaling λ_k's, μ_k's, the scaled queue length processes (when scaling space) behave according to the fluid dynamics. We note here that Atar *et al.* [8, 9] show in fact that the $\tilde{c}\mu/\theta$ rule is asymptotically fluid optimal in a multi-server setting and assuming overload. Due to the multi-server setting, the authors of

Atar *et al.* [8, 9] need a different limiting regime: the arrival rates and the number of servers are scaled, while the service rate of each server is kept fixed to μ. We do expect though, in the case of overload, that a similar proof technique can be applied to our model.

5.4 Heuristics for an arbitrary number of classes

In this section we will propose a heuristic for the stochastic optimization model with abandonments. This heuristic is based on the insights we obtained from the fluid control model.

We first consider the overload setting. In that case, the optimal fluid policy is to give priority according to the $\tilde{c}\mu/\theta$-rule. In Section 5.5 we will evaluate this policy when employed in the stochastic model (in overload).

We now consider the case of underload. Recall that in Proposition 5.1 we have seen that the optimal fluid control has a remarkable structure in the case of two classes: close to the origin the $\tilde{c}\mu$-rule is optimal, and when one of the fluids is sufficiently large the $\tilde{c}\mu/\theta$-rule is optimal. We observe the same structural property in the optimal solution for the stochastic control problem obtained numerically, see Section 5.3.3 and Figure 5.5 left). Our approach is thus to develop a heuristic that follows this insight, that is, close to the origin it will behave according to the $\tilde{c}\mu$-rule, and when the number of users in one of the classes is sufficiently large it will follow the $\tilde{c}\mu/\theta$-rule. It is not clear what should be the best choice for the threshold to decide whether the $\tilde{c}\mu$ rule or the $\tilde{c}\mu/\theta$ rule should be applied.

We propose the following heuristic, which is based on the two-class fluid analysis: For a general K-class queue we compare all classes pairwise and calculate the switching curves of the paired systems, see for example Figure 5.4 where $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2 \geq \tilde{c}_3\mu_3$ and $\tilde{c}_3\mu_3/\theta_3 \geq \tilde{c}_1\mu_1/\theta_1 \geq \tilde{c}_2\mu_2/\theta_2$. Then, whenever the state $(N_1, N_2, \ldots, N_k)$ satisfies that all pairs (N_i, N_j) lie under their corresponding switching curves, we give priority to the class with the highest value for $\tilde{c}\mu$. However, if there is at least one state (N_i, N_j) that lies above its corresponding switching curve, we will give priority to the class with the highest value for $\tilde{c}\mu/\theta$. For example, for the parameters of Figure 5.4 we will give priority to class 1 when both states (N_2, N_3) and (N_1, N_3) lie below their corresponding switching curves and otherwise priority is given to class 3. Whenever the queue of one class is empty we analyze the system in the same way but only take into account the $K-1$ queues that are non-empty. For a better understanding we give a pseudo-code of the heuristic rule in Algorithm 1.

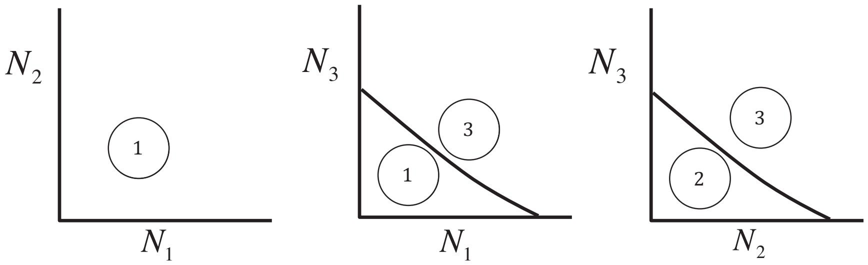


Figure 5.4: An example of the heuristics for the case K=3 when $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2 \geq \tilde{c}_3\mu_3$ and $\tilde{c}_3\mu_3/\theta_3 \geq \tilde{c}_1\mu_1/\theta_1 \geq \tilde{c}_2\mu_2/\theta_2$.

Algorithm 1 Algorithm to compute heuristic scheduling rule for an arbitrary K

Assume r queues are non empty.
Let N_i be the state of class $i \in \{1, \ldots, r\}$.
Compute the indices $\tilde{c}\mu$ and $\tilde{c}\mu/\theta$ for $i \in \{1, \ldots, r\}$.
Given a pair of classes i and j, such that $\tilde{c}_i\mu_i/\theta_i \geq \tilde{c}_j\mu_j/\theta_j$, compute the switching curve h_{ij} as given by Equation (5.3.1).
if for all $i, j, N_i \leq h_{ij}(N_j)$ **then**
    Give priority to the class with highest index $\tilde{c}\mu$
else
    Give priority to the class with highest index $\tilde{c}\mu/\theta$
end if

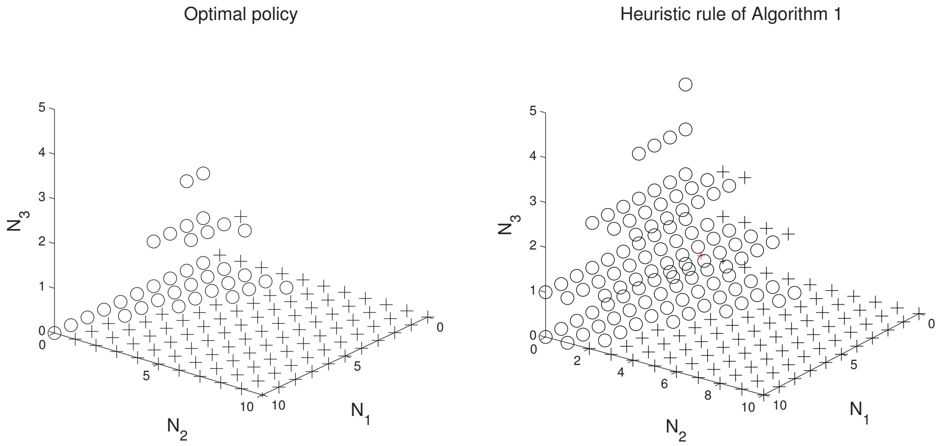


Figure 5.5: Optimal policy and heuristic for a 3-class single server example for the $\theta' = 0$ case. Circles indicate that class 1 is served, pluses that class 2 is served and absence of a sign corresponds to class 3 being served.

We propose an example with $K = 3$ to illustrate the heuristic we have just defined and to compare it to the optimal policy (obtained numerically by value iteration, see Section 1.3.3). Let us consider the following set of parameters $\mu = [10, 10, 9]; \theta = [1, 0.5, 0.25]; c = [1.7, 1.7, 1.7]; \delta = [2, 2, 4]; \lambda = [2, 2, 1]$. Hence, $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2 \geq \tilde{c}_3\mu_3$ and $\tilde{c}_1\mu_1/\theta_1 \leq \tilde{c}_2\mu_2/\theta_2 \leq \tilde{c}_3\mu_3/\theta_3$. Under our heuristic, class 1 will be served when all three classes are close enough to the origin, according to the $\tilde{c}\mu$ rule, and class 3 will be served otherwise, according to the $\tilde{c}\mu/\theta$ rule. Class 2 will be served in the following two cases: (i) when class 1 is empty and (N_2, N_3) is sufficiently close to the origin (follows from the $\tilde{c}\mu$ rule) and (ii) when class 3 is empty and (N_1, N_2) is sufficiently far from the origin (follows from the $\tilde{c}\mu/\theta$ rule). In Figure 5.5 we plot the actions under the optimal scheduling rule (calculated by value iteration) (left) and under our heuristic (right). We observe that the heuristic rule shows a qualitative similar structure to the optimal solution. In Section 5.5.2 we will present a numerical comparison of its performance.

5.5 Numerical results

In this section we numerically evaluate the performance of the heuristic described in Algorithm 1. We compare the performance of our heuristic rule against the optimal policy. The latter is calculated using the value iteration algorithm, introduced in Section 1.3.3. We also evaluate the following index policies:

- The $\tilde{c}\mu/\theta$-rule. This rule was introduced in Atar *et al.* [8, 9] where it was proved to be asymptotically fluid optimal for a multi-server system in overload. As shown in Proposition 5.2 this rule is also optimal for our fluid model in overload.
- The $\tilde{c}\mu/\theta - c$-rule. This rule was derived in Ayesta *et al.* [15] for the system $\theta' = 0$ (without arrivals) with the modification that the user in service also contributes to the cost.
- The $\tilde{c}\mu$-rule. This is the greedy or myopic rule that minimizes the instantaneous cost. This rule can be seen as a counterpart of the well-known $c\mu$-rule (Buyukkoc *et al.* [33]) for the system with abandonments.

The index rules $\tilde{c}\mu/\theta$ and $\tilde{c}\mu/\theta - c$ correspond to the Whittle index policy obtained in Proposition 3.3. The $\tilde{c}\mu$ rule corresponds to the Whittle index policy obtained in Proposition 3.7 for the particular case of linear holding cost.

Before we start describing in detail the results, we provide below our main conclusions:

- The qualitative performance in the $\theta = \theta'$ and $\theta' = 0$ cases are very similar.
- The $\tilde{c}\mu/\theta$ and the $\tilde{c}\mu/\theta - c$-rules perform very well in overload.
- Our heuristic (as proposed in Algorithm 1) performs very well across all loads.

For the sake of fairness we can mention that even though the index rules $\tilde{c}\mu/\theta$ and $\tilde{c}\mu$ perform worse than the heuristic rule, they are simpler to implement since they are state independent.

We now present the scenarios we have evaluated. In Section 5.5.1 we consider the case $K = 2$ and in Section 5.5.2 the case $K = 3$.

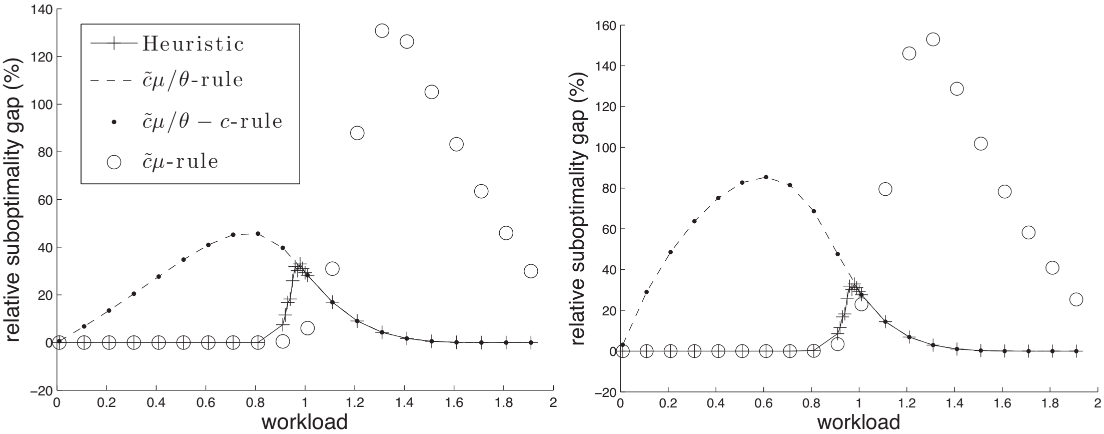
5.5.1 Performance analysis for two classes of customers

We consider the two models, $\theta = \theta'$ and $\theta' = 0$, and we calculate the relative suboptimality gap for the policies described above. In Examples 1 and 2 we fix the parameters c, δ, μ and θ and set $\rho_1 = \rho_2$ and vary the total workload ρ. In Example 3 we fix ρ and vary the value of θ_1.

Example 1: In this first example we set $\theta = [2, 0.05]; \mu = [14, 16]; c = [1, 1]; \delta = [4, 0.3]$, such that $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ and $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$. The results for this example are depicted in Figure 5.6a for the $\theta = \theta'$ case and in Figure 5.6b for the $\theta' = 0$ case.

In underload the heuristic and the $\tilde{c}\mu$-rule behave optimally, while the $\tilde{c}\mu/\theta$ and $\tilde{c}\mu/\theta - c$-rules behave very poorly. In Figure 5.3b we plotted the switching curves corresponding to the load $\rho = 0.73$. In fact, the average number of users (in the $\theta = \theta'$ system with $\rho = 0.73$) is given by $(\overline{N}_1, \overline{N}_2) = (0.4047, 1.5092)$ which is a state far below both the $\theta = \theta'$ switching curve and the fluid switching curve. This shows why both our heuristic and the $\tilde{c}\mu$-rule (this is the control below the switching curve) behave close to optimal.

In the overload case though, the $\tilde{c}\mu$-rule incurs a high relative suboptimality gap while our heuristic and the $\tilde{c}\mu/\theta$ and $\tilde{c}\mu/\theta - c$-rules are close to optimal. The latter is in agreement with what we expected as described in Section 5.3.3.



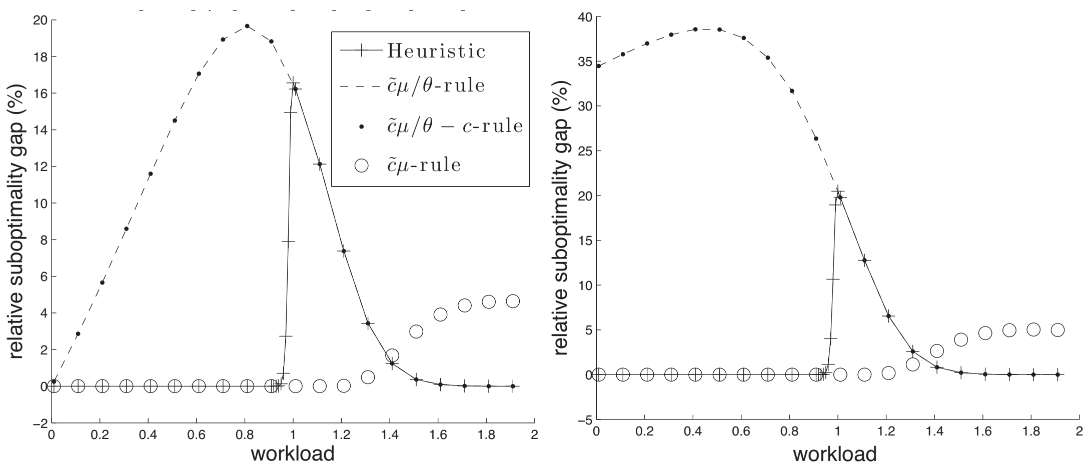
(a) Relative suboptimality gap for the $\theta = \theta'$ case (b) Relative suboptimality gap for the $\theta' = 0$ case

Figure 5.6: Performance comparison of policies for Example 1.

We observe that when the load is close to the critical regime $\rho = 1$ the suboptimality gap is around 30%. Our heuristic will give priority to class 2 in case $\rho > 1$ and has a switching curve very close to the origin in case $\rho = 1 - \epsilon$. In Figure 5.3c, which corresponds to the current example, we see that the optimal policy for the stochastic optimization problem is described by a switching curve for $\rho = 1$. Hence, when we are in a state below the switching curve, class 1 will be given priority. The process when $\rho = 1$ lives on average close to the stochastic switching curve, therefore, our policy can be far from optimal, as discussed in Section 5.3.3.

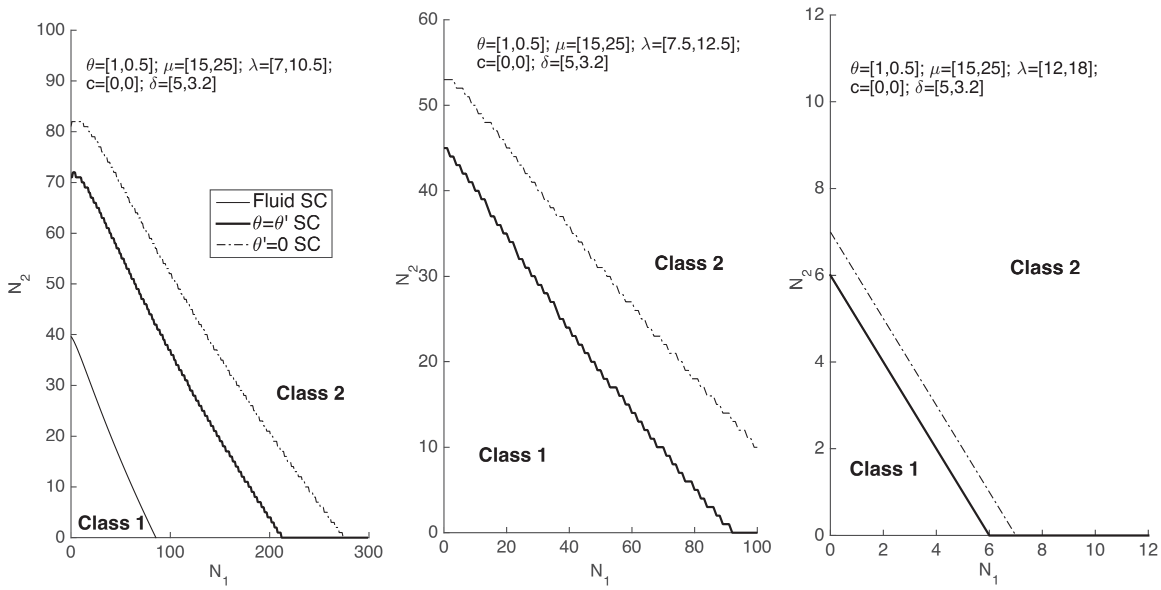
Example 2: In this second example we set $\theta = [1, 0.5]; \mu = [15, 25]; c = [0, 0]; \delta = [5, 3.2]$, so that $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ and $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$. As explained in Remark 5.3, setting $c_1 = c_2 = 0$ gives a different interpretation of the model: customers will abandon the system when a certain deadline is met before they have attained full service. In this case the $\tilde{c}\mu/\theta$, $\tilde{c}\mu/\theta - c$ and the $\tilde{c}\mu$ rules reduce to the $\delta\mu$, $\delta\mu$, and $\delta\theta\mu$ rules, respectively. We observe that the index $\delta\mu$ does not perform well in underload but is close to optimal in overload. The opposite holds for the $\delta\theta\mu$ rule. Our policy is optimal in underload and as good as the $\delta\mu$ index in overload, see Figure 5.7a for the $\theta = \theta'$ case and in Figure 5.7b for the $\theta' = 0$ case.

In Figure 5.8 we plotted the optimal switching curves for the stochastic models $\theta = \theta'$ and $\theta' = 0$ (obtained by value iteration), as well as the optimal fluid switching curve $h(\cdot)$. Figure 5.8a corresponds to load $\rho = 0.8867$. In that case, the average number of customers is given by $(\overline{N}_1, \overline{N}_2) = (0.6859, 2.6963)$, which is a state far below all switching curves. Hence, this shows why our heuristic and the $\tilde{c}\mu$-rule perform close to optimal. On the other hand, when $\rho = 1$, Figure 5.8b, the optimal control in the fluid model is to serve class 2, so there is no switching curve. Under the optimal policy for the stochastic model, the average number of customers is given by $(\overline{N}_1, \overline{N}_2) = (0.763, 4.2703)$. This is a state far below the switching curves of the stochastic model. Hence, most of the time priority is given to class 1 under the optimal policy. This explains why our heuristic gives a positive optimality gap of 16%. However, as the load of the system increases ($\rho > 1$) the process will live more above the optimal switching curve for the stochastic model. See for example Figure 5.8c for load $\rho = 1.52$ for which the average number of customers under the optimal policy is given by $(\overline{N}_1, \overline{N}_2) = (6.8054, 3.7244)$. This explains why our heuristic, which gives priority to class 2, has a suboptimality gap very close to 0%.



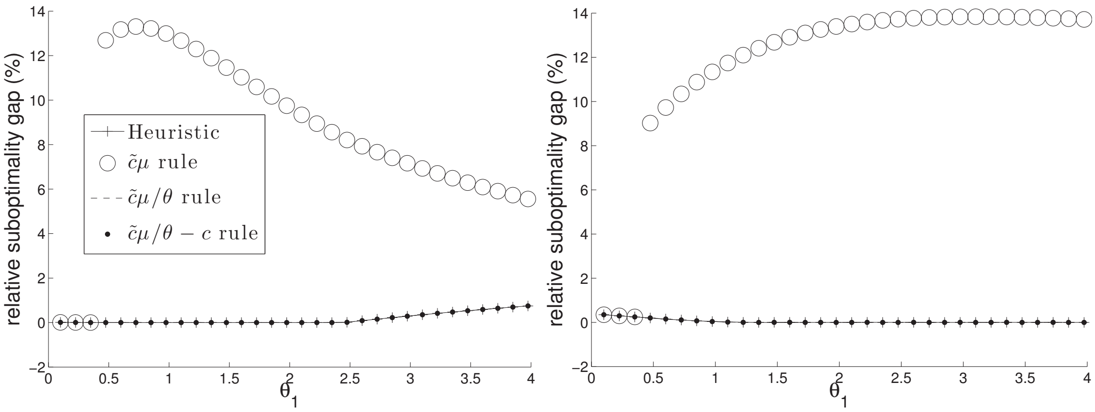
(a) Relative suboptimality gap for the $\theta = \theta'$ case (b) Relative suboptimality gap for the $\theta' = 0$ case

Figure 5.7: Performance comparison of policies for Example 2.



(a) $\rho = 0.8867$. (b) $\rho = 1$. (c) $\rho = 1.52$.

Figure 5.8: Comparison of switching curves for Example 2.



(a) Relative suboptimality gap for the $\theta = \theta'$ case (b) Relative suboptimality gap for the $\theta' = 0$ case

Figure 5.9: Performance comparison of policies for Example 3, with load $\rho = 0.7$.

Remark 5.4 (Peak when workload close to 1)**.** *We observe in Examples 1 and 2 that when the workload is close to 1 a peak appears in the suboptimality gap for the heuristic rule. This can be explained by the following. In the proof of Proposition 5.1, see Appendix 5.6.2, we observe a switching curve whenever*

$$h(0) = (1 - \rho_1 - \rho_2)\frac{\mu_2}{\theta_1\theta_2}\left(\frac{\tilde{c}_1\mu_1 - c_2\tilde{\mu}_2}{\dfrac{\tilde{c}_2\mu_2}{\theta_2} - \dfrac{\tilde{c}_1\mu_1}{\theta_1}}\right) > 0.$$

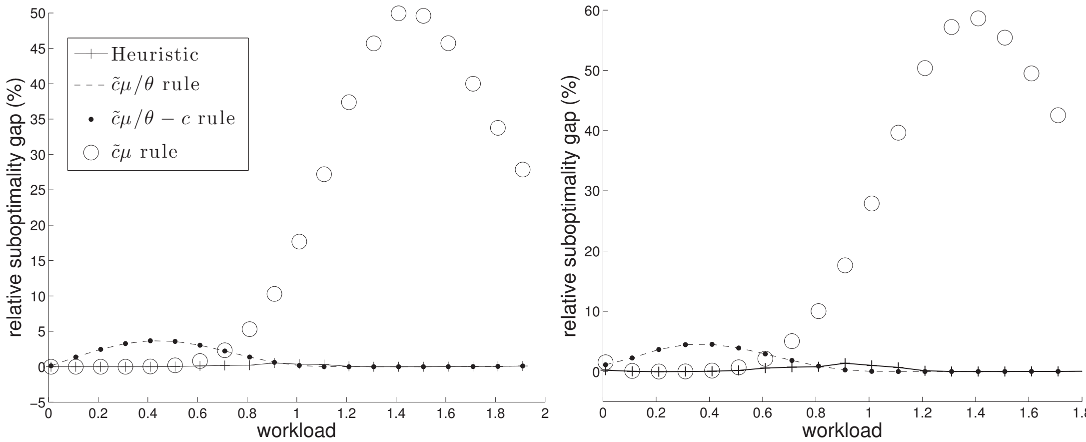
Therefore, as $1 - \rho_1 - \rho_2 \to 0$ the switching curve vanishes, that is, the heuristic becomes equivalent to the $\tilde{c}\mu/\theta$-rule. However, around $\rho = 1$ the optimal stochastic control still follows the $\tilde{c}\mu$-rule in a non-negligible part of the state space, see for instance Figure 5.8b.

Example 3: We consider the following parameters: $\theta_2 = 0.1; \mu = [8, 8]; \lambda = [2.8, 2.8]; c = [1, 1]; \delta = [0.5, 2]$, and we let θ_1 vary. Hence, $\rho = 0.7$, i.e., we are in underload. The results are plotted in Figure 5.9a for the $\theta = \theta'$ case and in Figure 5.9b for the $\theta' = 0$ case.

When $\theta_1 \in [0, 0.4]$, we have $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$ and $\tilde{c}_2\mu_2 \geq \tilde{c}_1\mu_1$, in which case the heuristic gives priority to class 2, as do all the index policies. On the other hand, when $\theta_1 \in (0.4, 4]$, then $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$ and $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ and a switching curve does appear in the heuristic. For the cases where no switching curve appears ($\theta_1 \in [0, 0.4]$), all the index rules are optimal, but as soon as a switching curve emerges in the heuristic the $\tilde{c}\mu$ rule gives a positive suboptimality gap. The reason why the $\tilde{c}\mu$-rule performs bad in this particular case is that as soon as the ratio $\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\tilde{c}_2\mu_2/\theta_2 - \tilde{c}_1\mu_1/\theta_1}$ becomes small, the switching curve get close to zero and hence the $\tilde{c}\mu/\theta$ rule becomes optimal.

5.5.2 Performance analysis for arbitrary number of customers

We analyze the relative performance of the heuristic as explained in Section 5.4. Here we take the same example that was introduced in Section 5.4 with parameters $\mu = [10, 10, 9]; \theta = [1, 0.5, 0.25]; c = [1.7, 1.7, 1.7]; \delta = [2, 2, 4]$. Let $\lambda_i = \lambda\beta_i$, $i = 1, 2, 3$, denote the arrival rate of class-i jobs, where λ denotes the total arrival rate and β_i is the fraction of class-i customers. We choose β_i, $i = 1, 2, 3$, in such a way that $\rho_1 = \rho_2 = \rho_3$. We vary the value of λ to change the total load in the system and we compute the relative



(a) Relative suboptimality gap for the $\theta = \theta'$ case (b) Relative suboptimality gap for $\theta' = 0$ case

Figure 5.10: Performance comparison of policies for $K = 3$.

suboptimality gap of all the policies, including the heuristic. The results are depicted in Figure 5.10. We observe that our policy is optimal together with the $\tilde{c}\mu$ rule for very low loads but at some point, as the workload ρ increases, the $\tilde{c}\mu$-rule starts performing increasingly worse. On the other hand, the $\tilde{c}\mu/\theta$ and the $\tilde{c}\mu/\theta - c$ rules become optimal when the load becomes larger than 1. Our heuristic keeps a low suboptimality gap throughout.

5.6 Appendix

5.6.1 Proof of Lemma 5.1

Lemma 5.1 states which class is optimal to serve close to the origin. We calculate the cost function when starting in a point very close to the origin $(m_1(0), m_2(0)) = (\varepsilon, \varepsilon)$ when priority is given to class 1. We do the same for the case when priority is given instead to class 2. When comparing both cost functions, we get the condition under which prioritizing class 1 gives lower cost than prioritizing class 2. We note that it is sufficient to compare the above described two policies: since the control appears linearly, we can assume that under the optimal policy full priority will be given to one class as long as we start close enough to the origin.

We first consider the control that gives full priority to class 1. When class 1 hits zero, ρ_1 is given to class 1 and $1 - \rho_1$ to class 2, until the equilibrium $(0, 0)$ is reached. The cost under this policy, starting in state $(m_1(0), m_2(0)) = (\varepsilon, \varepsilon)$, is

$$C_1(t, m) := \int_0^T \tilde{c}_1 m_1(t) + \tilde{c}_2 m_2(t) \, \mathrm{d}t.$$

In order to compute the trajectories $m_1(t)$ and $m_2(t)$, we will split up the time into two time intervals, $[0, t_1]$ and $[t_1, t_2]$, where t_1 is the moment when class 1 hits zero and t_2 when class 2 hits zero. After some algebra, we obtain that for the interval $[0, t_1]$ the trajectories are as follows:

$$\begin{cases} m_1(t) = \left(\varepsilon + \dfrac{\mu_1 - \lambda_1}{\theta_1}\right) \mathrm{e}^{-\theta_1 t} + \dfrac{\lambda_1 - \mu_1}{\theta_1} = -\theta_1 t \left(\varepsilon + \dfrac{\mu_1 - \lambda_1}{\theta_1}\right) + \varepsilon + o(\varepsilon), & t \in [0, t_1], \\ m_2(t) = \left(\varepsilon - \dfrac{\lambda_2}{\theta_2}\right) \mathrm{e}^{-\theta_2 t} + \dfrac{\lambda_2}{\theta_2} = \theta_2 t \left(\dfrac{\lambda_2}{\theta_2} - \varepsilon\right) + \varepsilon + o(\varepsilon), & t \in [0, t_1]. \end{cases}$$

We used here that $t_2 \leq \frac{m_1(0)/\mu_1 + m_2(0)/\mu_2}{1-\rho} = O(\varepsilon)$ [3], hence $\mathrm{e}^{-\theta_1 t} = -\theta_1 t + 1 + o(\varepsilon)$, for $t \leq t_2$. (Here $o(\varepsilon) = g(\varepsilon)$ for $g(\cdot)$ a function that satisfies $\lim_{\varepsilon \to 0} g(\varepsilon)/\varepsilon = 0$.) We note that since ε is chosen small enough, $m_2(t) > 0$ for all $t < t_1$. Time t_1 being the moment at which class 1 empties, we obtain

$$t_1 = \frac{\varepsilon}{\theta_1 \left(\varepsilon + \dfrac{\mu_1 - \lambda_1}{\theta_1}\right)} = \frac{\varepsilon}{\mu_1 - \lambda_1} + o(\varepsilon),$$

therefore

$$m_2(t_1) = \frac{\lambda_2}{\theta_2} \frac{\varepsilon \theta_2}{\mu_1 - \lambda_1} + \varepsilon + o(\varepsilon). \tag{5.6.1}$$

Recall that t_2 is the time at which class 2 is emptied. In the interval $[t_1, t_2]$ class 1 receives service ρ_1 and class 2 service $1 - \rho_1$. Hence, after some algebra we obtain that

$$\begin{cases} m_1(t) = 0, & t \in [t_1, t_2], \\ m_2(t) = A_2' \mathrm{e}^{-\theta_2 t} + \dfrac{\lambda_2 - \mu_2(1-\rho_1)}{\theta_2} = A_2' \left(-\theta_2 t + 1\right) + \dfrac{\lambda_2 - \mu_2(1-\rho_1)}{\theta_2} + o(\varepsilon), & t \in [t_1, t_2], \end{cases}$$

where A_2' is the constant of integration. Here we used that $t = O(\varepsilon)$, hence $\mathrm{e}^{-\theta_2 t} = -\theta_2 t + 1 + o(\varepsilon)$. Moreover, from (5.6.1) we obtain

$$A_2' = \frac{-\dfrac{\lambda_2}{\theta_2}\left(\dfrac{-\varepsilon\theta_2}{\mu_1 - \lambda_1}\right) + \varepsilon + \dfrac{\mu_2(1-\rho_1) - \lambda_2}{\theta_2}}{1 - \dfrac{\theta_2 \varepsilon}{\mu_1 - \lambda_1}} + o(\varepsilon),$$

hence, we have

$$\begin{aligned} m_2(t) &= \frac{(\lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1))\varepsilon + (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)}{-\theta_2^2 \varepsilon + (\mu_1 - \lambda_1)\theta_2} (-\theta_2 t + 1) + \frac{\lambda_2 - \mu_2(1-\rho_1)}{\theta_2} + o(\varepsilon), \\ &= \frac{(\lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1))\varepsilon + (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)}{-\theta_2^2 \varepsilon + (\mu_1 - \lambda_1)\theta_2} (-\theta_2 t) \\ &\quad + \frac{(\mu_1 - \lambda_1 + \mu_2(1-\rho_1))\varepsilon}{-\theta_2 \varepsilon + \mu_1 - \lambda_1} + o(\varepsilon), \qquad t \in [t_1, t_2], \end{aligned}$$

[3] This follows from the fact that the workload $w(t) := m_1(t)/\mu_1 + m_2(t)/\mu_2$ has a negative drift smaller than or equal to $\rho - 1$, see the footnote in Section 5.3.

and from $m_2(t_2) = 0$ we obtain

$$
\begin{aligned}
t_2 &= \frac{(\mu_1 - \lambda_1 + \mu_2(1-\rho_1))\varepsilon}{-\theta_2^2\varepsilon^2 + (\lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1))\varepsilon + (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)} + o(\varepsilon) \\
&= \frac{(\mu_1 - \lambda_1 + \mu_2(1-\rho_1))\varepsilon}{(\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)} + o(\varepsilon).
\end{aligned}
$$

We can now compute the cost function:

$$
\begin{aligned}
C_1(t,(\varepsilon,\varepsilon)) \quad &= \quad \int_0^{t_1} \tilde{c}_1 \left(\left(\varepsilon + \frac{\mu_1 - \lambda_1}{\theta_1} \right) (-\theta_1 t) + \varepsilon \right) + \tilde{c}_2 \left(\left(\varepsilon - \frac{\lambda_2}{\theta_2} \right) (-\theta_2 t) + \varepsilon \right) \,\mathrm{d}t \\
&\quad + \int_{t_1}^{t_2} \tilde{c}_2 \left(\frac{(\lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1))\varepsilon + (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)}{-\theta_2^2\varepsilon + (\mu_1 - \lambda_1)\theta_2} (-\theta_2 t) \right) \,\mathrm{d}t \\
&\quad + \int_{t_1}^{t_2} \tilde{c}_2 \left(\frac{(\mu_1 - \lambda_1 + \mu_2(1-\rho_1))\varepsilon}{-\theta_2\varepsilon + \mu_1 - \lambda_1} \right) \,\mathrm{d}t + o(\varepsilon^2) \\
&= \quad \varepsilon^2 \left(\frac{\tilde{c}_1}{2(\mu_1 - \lambda_1)} + \tilde{c}_2 \frac{2(\mu_1 - \lambda_1) + \lambda_2}{2(\mu_1 - \lambda_1)^2} \right) \\
&\quad - \tilde{c}_2 \frac{(\lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1))\varepsilon + (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)}{2(-\theta_2\varepsilon + (\mu_1 - \lambda_1))} \left((t_2)^2 - (t_1)^2 \right) \\
&\quad + \tilde{c}_2 \left(\frac{(\mu_1 - \lambda_1 + \mu_2(1-\rho_1))\varepsilon}{-\theta_2\varepsilon + \mu_1 - \lambda_1} \right) (t_2 - t_1) + o(\varepsilon^2),
\end{aligned}
$$

where

$$
(t_2)^2 - (t_1)^2 = \frac{(b_1\varepsilon^2 + b_2\varepsilon)^2}{(-b_1^2\varepsilon^2 + b_3\varepsilon + b_4)^2} - \varepsilon^2 b_5^2 + o(\varepsilon^2) = \frac{b_2^2\varepsilon^2}{b_4^2} - \varepsilon^2 b_5^2 + o(\varepsilon^2),
$$

$$
t_2 - t_1 = \frac{b_1\varepsilon^2 + b_2\varepsilon}{b_4} - b_5\varepsilon + o(\varepsilon) = \left(\frac{b_2}{b_4} - b_5 \right) \varepsilon + o(\varepsilon^2) + o(\varepsilon),
$$

with

$$
b_1 = -\theta_2,\ b_2 = \mu_1 - \lambda_1 + \mu_2(1-\rho_1),
$$

$$
b_3 = \lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1),\ b_4 = (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1),\ b_5 = \frac{1}{\mu_1 - \lambda_1}.
$$

After some calculations, we then obtain

$$
C_1(t,(\varepsilon,\varepsilon)) = \tilde{c}_1\varepsilon^2 \left(\frac{1}{2(\mu_1 - \lambda_1)} \right) + \tilde{c}_2\varepsilon^2 \left(\frac{2(\mu_1 - \lambda_1) + \lambda_2}{2(\mu_1 - \lambda_1)^2} + \frac{(\mu_1 - \lambda_1 + \lambda_2)^2}{2(\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)^2} \right) + o(\varepsilon^2).
$$

By symmetry, the cost when instead class 2 is given priority is given by

$$
C_2(t,(\varepsilon,\varepsilon)) = \tilde{c}_2\varepsilon^2 \left(\frac{1}{2(\mu_2 - \lambda_2)} \right) + \tilde{c}_1\varepsilon^2 \left(\frac{2(\mu_2 - \lambda_2) + \lambda_1}{2(\mu_2 - \lambda_2)^2} + \frac{(\mu_2 - \lambda_2 + \lambda_1)^2}{2(\mu_1(1-\rho_2) - \lambda_1)(\mu_2 - \lambda_2)^2} \right) + o(\varepsilon^2).
$$

It can now be checked that $C_1(t,(\varepsilon,\varepsilon)) \leq C_2(t,(\varepsilon,\varepsilon))$ if and only if $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$, (given that we are in underload ($\rho < 1$)), which proves the result.

5.6.2 Proof of Proposition 5.1

The proof of this proposition follows from the PMP, *i.e.*, the necessary conditions for optimality explained in Appendix A.2, and Lemma 5.1. We will therefore refer to them throughout the proof.

We begin by proving that the qualification conditions, as given in Appendix A.2, hold in order to be able to apply PMP. Note that $h_1(s(t)) = (-1 + s_1(t) + s_2(t), -s_1(t), -s_2(t))^\mathsf{T}$ and $h_2^1(m(t)) = (-\lambda_1 + \mu_1 s_1(t) + \theta_1 m_1(t), -\lambda_2 + \mu_2 s_2(t) + \theta_2 m_2(t))^\mathsf{T}$ and therefore

$$\text{rank}\begin{bmatrix} \frac{\partial h_1(s(t))}{\partial s(t)} & \text{diag}(h_1) \end{bmatrix} = \text{rank}\begin{bmatrix} 1 & 1 & -1+s_1(t)+s_2(t) & 0 & 0 \\ -1 & 0 & 0 & -s_1(t) & 0 \\ 0 & -1 & 0 & 0 & -s_2(t) \end{bmatrix} = 3,$$

and

$$\text{rank}\begin{bmatrix} \frac{\partial h_2^1(m(t))}{\partial s(t)} \end{bmatrix} = \text{rank}\begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} = 2.$$

Having proven the constraint qualification to be satisfied, the proof reads as follows. We apply the PMP to obtain extremal solutions, that is, solutions that satisfy necessary conditions for optimality. We obtain four possible candidates: if $\tilde{c}\mu$ and $\tilde{c}\mu/\theta$ have the same ordering (*i.e.,* $\tilde{c}_i\mu_i \geq \tilde{c}_j\mu_j$ and $\tilde{c}_i\mu_i/\theta_i \geq \tilde{c}_j\mu_j/\theta_j$ with $i \neq j \in \{1,2\}$) then a strict priority rule is optimal and if $\tilde{c}\mu$ and $\tilde{c}\mu/\theta$ have opposite ordering (*i.e.,* $\tilde{c}_i\mu_i \geq \tilde{c}_j\mu_j$ and $\tilde{c}_j\mu_j/\theta_j \geq \tilde{c}_i\mu_i/\theta_i$ with $i \neq j \in \{1,2\}$), then a switching curve emerges, where above it one class is prioritized and below it the other class. Therefore, given the ordering of $\tilde{c}\mu$ and $\tilde{c}\mu/\theta$ one encounters two extremal solutions. Then, Lemma 5.1 serves to compare both candidates and chooses the optimal solution.

By assumption $\rho_1 + \rho_2 < 1$ and as explained in Section 5.3 under any non-idling control the optimal final time is finite and free (subject to optimization). We now write the Hamiltonian and the Lagrangian of Theorem A.2 with respect to our particular problem, that is, the Hamiltonian of the system writes

$$\mathcal{H}(m(t), s(t), \gamma(t)) = \sum_{k=1}^{2} (\tilde{c}_k m_k(t) + \gamma_k(t)(\lambda_k - \mu_k s_k(t)) - \theta_k m_k(t))),$$

and the Lagrangian

$$\begin{aligned}\mathcal{L}(m(t), s(t), \gamma(t), \nu(t), \omega(t)) =& \mathcal{H}(m(t), s(t), \gamma(t)) - \nu_1(t)m_1(t) - \nu_2(t)m_2(t) \\ & - \omega_1(t)s_1(t) - \omega_2(t)s_2(t) + \omega_3(t)(s_1(t) + s_2(t) - 1),\end{aligned}$$

where $\gamma_k(\cdot)$ is the adjoint variable for class-k fluid, $\nu_i(\cdot)$ for $i = 1, 2$ are the Lagrange multipliers that correspond to the state constraints, and $\omega_i(\cdot)$ for $i = 1, 2, 3$ are the Lagrange multipliers for the constraints on the control variables.

We first solve for Equation (A.2.3) for all time intervals for which $m_1(t), m_2(t) > 0$ (we will call this intervals *interior arcs*), which gives us $\gamma_k^*(t) = C_k' \mathrm{e}^{\theta_k t} + \frac{\tilde{c}_k}{\theta_k}$ for $k = 1, 2$, where C_k' are constants of integration. Also from (A.2.7), we have $\nu_1(t) = \nu_2(t) = 0$. From (A.2.5) and (A.2.7) we obtain $-\gamma_1(t)\mu_1 + \omega_3(t) = -\gamma_2(t)\mu_2 + \omega_3(t) = 0 \Rightarrow \mu_1\gamma_1(t) = \mu_2\gamma_2(t)$, for all time intervals for which either $m_1(t) = 0, m_2(t) > 0$ or $m_2(t) = 0, m_1(t) > 0$, we will call this intervals *boundary arcs.* Note also that in a boundary arc with $m_k(t) = 0$, $\mathrm{d}m_k(t)/\mathrm{d}t = 0$ and therefore $s_k(t) = \lambda_k/\mu_k$ and $\omega_k(t) = 0$ for $k = 1, 2$.

We now solve for Equation (A.2.4) in interior arcs, which is equivalent to solving

$$\arg\min_{\mathbf{s}\in\mathcal{S}} \sum_{k=1}^{2} -\mu_k\gamma_k(t)s_k(t) = \arg\min_{\mathbf{s}\in\mathcal{S}} \sum_{k=1}^{2} -\mu_k \left(C_k'\mathrm{e}^{\theta_k t} + \frac{\tilde{c}_k}{\theta_k}\right) s_k(t), \tag{5.6.2}$$

The latter implies that under the optimal control the class with highest value for $\mu_k \left(C_k'\mathrm{e}^{\theta_k t} + \frac{\tilde{c}_k}{\theta_k}\right)$ will be prioritized in interior arcs. Once the constants of integration C_k' for $k = 1, 2$ are known one can completely characterize the adjoint vectors and the Lagrange multipliers.

We will therefore compute the constants C_k' for $k = 1, 2$, which will help in establishing the priority rules. First, to determine whether there is a switch in priorities we study the following switching function:

$$\sigma(t) := \mu_1 \left(C_1'\mathrm{e}^{\theta_1 t} + \frac{\tilde{c}_1}{\theta_1}\right) - \mu_2 \left(C_2'\mathrm{e}^{\theta_2 t} + \frac{\tilde{c}_2}{\theta_2}\right).$$

Observe that function $\sigma(t)$ has at most two roots, the extremal solution will therefore be one out of this three options:

(i) It is a strict priority rule, that is, $\nexists\, t \in [0, T]$ s.t $\sigma(t) = 0$.

(ii) It is to prioritize one of the classes close to the origin and the other far from the origin, that is, $\exists!\, t_1 \in [0, T],$ s.t $\sigma(t_1) = 0$.

(iii) It is to give priority to one of the classes close to the origin and far from it and to the other class in between, that is, $\exists\, t_1, t_2 \in [0, T]$, such that $t_1 > t_2$ and $\sigma(t_1) = \sigma(t_2) = 0$.

Let us assume w.l.o.g $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$. The other case can be analyzed using the same arguments but with the indices swapped. We will apply PMP to show that (iii) is never an optimal solution, and we will derive the conditions under which (i) or (ii) satisfy the necessary conditions.

To do so, let us first assume that there exists at least one switching curve. For ease of notation we assume the initial state $(m_1(0), m_2(0)) = (m_{10}, m_{20})$, $m_{10}, m_{20} > 0$ to be a point on the switching curve. In the case of two switching curves we assume that (m_{10}, m_{20}) is on the last switching curve before reaching the equilibrium point $(0, 0)$. Recall now the assumption $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$, therefore by Lemma 5.1 we will have that $s^*(t) = (0, 1)$ for $t_1 = 0$, $s^*(t) = (1, 0)$ for $t \in (t_1, t_2]$, where t_2 is the time at which the trajectory enters the boundary arc $m_1^*(t) = 0$ and $s^*(\rho_1, 1 - \rho_1)$ for all $t \in (t_2, t_3]$, where t_3 is the time at which equilibrium is reached.

We will now study the switching function, which fully characterizes which class is given priority. In order to obtain C_1', C_2', we will apply the transversality conditions of the PMP in Equation (A.2.8).

If $t = t_1 = 0$, then $s_2^*(t) = 1$ and $s_1^*(t) = 0$. Using that $\gamma_k^*(t) = C_k'\mathrm{e}^{\theta_k t} + \tilde{c}_k/\theta_k$, we obtain that the Hamiltonian for $t = t_1 = 0$ is given by

$$\mathcal{H}(m^*(t), s^*(t), \gamma^*(t), t) = \tilde{c}_1\frac{\lambda_1}{\theta_1} - C_1'\theta_1\left(m_{10} - \frac{\lambda_1}{\theta_1}\right) + \tilde{c}_2\left(\frac{\lambda_2 - \mu_2}{\theta_2}\right) - \theta_2 C_2'\left(m_{20} + \frac{\mu_2 - \lambda_2}{\theta_2}\right). \tag{5.6.3}$$

If $t \in (0, t_2]$, then $s_1^*(t) = 1$ and $s_2^*(t) = 0$. Hence,

$$m_1^*(t) = \left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1}\right)\mathrm{e}^{-\theta_1 t} + \frac{\lambda_1 - \mu_1}{\theta_1}, \qquad m_2^*(t) = \left(m_{20} - \frac{\lambda_2}{\theta_2}\right)\mathrm{e}^{-\theta_2 t} + \frac{\lambda_2}{\theta_2},$$

so that for $t \in (0, t_2]$ we have

$$\begin{aligned}
&\mathcal{H}(m^*(t), s^*(t), \gamma^*(t), t)\\
&= \tilde{c}_1\left(\left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1}\right)\mathrm{e}^{-\theta_1 t} + \frac{\lambda_1 - \mu_1}{\theta_1}\right) + \tilde{c}_2\left(\left(m_{20} - \frac{\lambda_2}{\theta_2}\right)\mathrm{e}^{-\theta_2 t} + \frac{\lambda_2}{\theta_2}\right)\\
&\quad + \left(C_1'\mathrm{e}^{\theta_1 t} + \frac{\tilde{c}_1}{\theta_1}\right)\left(\lambda_1 - \mu_1 - \theta_1\left(\left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1}\right)\mathrm{e}^{-\theta_1 t} + \frac{\lambda_1 - \mu_1}{\theta_1}\right)\right)\\
&\quad + \left(C_2'\mathrm{e}^{\theta_2 t} + \frac{\tilde{c}_2}{\theta_2}\right)\left(\lambda_2 - \theta_2\left(\left(m_{20} - \frac{\lambda_2}{\theta_2}\right)\mathrm{e}^{-\theta_2 t} + \frac{\lambda_2}{\theta_2}\right)\right)\\
&= \tilde{c}_1\left(\frac{\lambda_1 - \mu_1}{\theta_1}\right) + \tilde{c}_2\frac{\lambda_2}{\theta_2} - \theta_1 C_1'\left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1}\right) - \theta_2 C_2'\left(m_{20} - \frac{\lambda_2}{\theta_2}\right). \qquad (5.6.4)
\end{aligned}$$

Setting (5.6.3) and (5.6.4) equal to 0, we obtain the following expressions:

$$C_1' = \frac{\tilde{c}_1\left(\frac{\lambda_1 - \mu_1}{\theta_1}\right) + \tilde{c}_2\frac{\lambda_2}{\theta_2} - \theta_2 C_2'\left(m_{20} - \frac{\lambda_2}{\theta_2}\right)}{\theta_1\left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1}\right)},$$

$$C_2' = \frac{\left(\frac{\tilde{c}_1\mu_1}{\theta_1} - \frac{\tilde{c}_2\mu_2}{\theta_2}\right)\theta_1 m_{10} - \mu_1\frac{\tilde{c}_2\mu_2}{\theta_2}(1 - \rho_1 - \rho_2)}{\mu_1\theta_2 m_{20} + \theta_1\mu_2 m_{10} + \mu_1\mu_2(1 - \rho_1 - \rho_2)}. \qquad (5.6.5)$$

If $t \in (t_2, t_3]$, then $s_1^*(t) = \rho_1$, as computed above, and $s_2^*(t) = 1 - \rho_1$. Hence,

$$m_1^*(t) = 0,$$

$$m_2^*(t) = \left(m_{20} - \frac{\lambda_2}{\theta_2} + \frac{\mu_2\mu_1 - \lambda_1\mu_2}{\mu_1\theta_2}\left(\frac{\mu_1 - \lambda_1}{m_{10}\theta_1 - \lambda_1 + \mu_1}\right)^{-\frac{\theta_2}{\theta_1}}\right)\mathrm{e}^{-\theta_2 t} - \left(\frac{\mu_2\mu_1 - \lambda_1\mu_2 - \lambda_2\mu_1}{\mu_1\theta_2}\right),$$

so that for $t \in (t_2, t_3]$ we have

$$\begin{aligned}
\mathcal{H}(m^*(t), s^*(t), \gamma^*(t), t) &= \tilde{c}_2 m_2^*(t) + \left(C_2'\mathrm{e}^{\theta_2 t} + \frac{\tilde{c}_2}{\theta_2}\right)\left(\lambda_2 - \mu_2\left(1 - \frac{\lambda_1}{\mu_1}\right) - \theta_2 m_2^*(t)\right)\\
&= -\frac{\tilde{c}_2\mu_2}{\theta_2}(1 - \rho_1 - \rho_2)\\
&\quad -\theta_2 C_2'\left(m_{20} - \frac{\lambda_2}{\theta_2} + \frac{\mu_2\mu_1 - \lambda_1\mu_2}{\mu_1\theta_2}\left(\frac{\mu_1 - \lambda_1}{m_{10} - \lambda_1 + \mu_1}\right)^{-\frac{\theta_2}{\theta_1}}\right). \qquad (5.6.6)
\end{aligned}$$

Setting Equation (5.6.6) equal to 0, and substituting the expression of C_1' and C_2' as given by Equation (5.6.5), we obtain that a state on the switching curve (the closest to the origin) satisfies the following relation:

$$m_{20} = \frac{a_1 m_{10} + a_2 + (a_3 m_{10} - a_2)\left(\frac{\theta_1 m_{10} + \mu_1 - \lambda_1}{\mu_1 - \lambda_1}\right)^{\frac{\theta_2}{\theta_1}}}{a_4 m_{10}} + \frac{\lambda_2}{\theta_2}, \qquad (5.6.7)$$

where

$$a_1 = \tilde{c}_2 \frac{\mu_2}{\theta_2} (1 - \rho_1 - \rho_2)\,; \;\; a_2 = a_1 \frac{\mu_1}{\theta_1} (1 - \rho_1);$$

$$a_3 = \left(\tilde{c}_1 \frac{\mu_1}{\theta_1} - \tilde{c}_2 \frac{\mu_2}{\theta_2} \right) (1 - \rho_1); \;\; a_4 = - \left(\tilde{c}_1 \frac{\mu_1}{\theta_1} - \tilde{c}_2 \frac{\mu_2}{\theta_2} \right) \frac{\theta_2}{\mu_2}.$$

This latter expression gives a switching curve in the first quadrant provided that $m_{20} > 0$ as $m_{10} \to 0$ Using l'Hopital we obtain

$$m_{20} \xrightarrow{m_{10} \to 0} \frac{a_1}{a_4} + \frac{\lambda_2}{\theta_2} + \frac{a_3}{a_4} - \frac{a_2 \theta_2}{a_4 \mu_1 (1 - \rho_1)} = (1 - \rho_1 - \rho_2) \frac{\mu_2}{\theta_1 \theta_2} \left(\frac{\tilde{c}_1 \mu_1 - \tilde{c}_2 \mu_2}{\frac{\tilde{c}_2 \mu_2}{\theta_2} - \frac{\tilde{c}_1 \mu_1}{\theta_1}} \right).$$

Since we assumed that the system is in under-load ($\rho_1 + \rho_2 < 1$), and w.l.o.g we assumed $\tilde{c}_1 \mu_1 \geq \tilde{c}_2 \mu_2$ we have

$$(1 - \rho_1 - \rho_2) \frac{\mu_2}{\theta_1 \theta_2} \left(\frac{\tilde{c}_1 \mu_1 - \tilde{c}_2 \mu_2}{\frac{\tilde{c}_2 \mu_2}{\theta_2} - \frac{\tilde{c}_1 \mu_1}{\theta_1}} \right) \geq 0 \Longleftrightarrow \tilde{c}_2 \mu_2 / \theta_2 > \tilde{c}_1 \mu_1 / \theta_1 \text{ and } \tilde{c}_1 \mu_1 \geq \tilde{c}_2 \mu_2.$$

The same with the indices swapped can be obtained assuming $\tilde{c}_2 \mu_2 \geq \tilde{c}_1 \mu_1$. Therefore, the condition so that a switch of priority occurs (at least one) is that the $\tilde{c}\mu$ and the $\tilde{c}\mu/\theta$ have the opposite ordering. We have hence proven that if $\tilde{c}\mu/\theta$ and $\tilde{c}\mu$ have the same ordering, there is no switch at all. The latter together with Lemma 5.1 implies that priority will be given to the class of customers with highest index $\tilde{c}\mu$ (or equivalently highest $\tilde{c}\mu/\theta$).

Having obtained the control in the case in which $\tilde{c}\mu/\theta$ and $\tilde{c}\mu$ have the same ordering, we are left with the case in which $\tilde{c}\mu/\theta$ and $\tilde{c}\mu$ have opposite ordering. In the latter case (ii) and (iii) could hold. We will show next that (iii) will never happen.

Let us assume $\tilde{c}_2 \mu_2 / \theta_2 > \tilde{c}_1 \mu_1 / \theta_1$ and $\tilde{c}_1 \mu_1 \geq \tilde{c}_2 \mu_2$. We will assume that there exist two switching curves and we will reach a contradiction. Let the initial point be on the switching curve that is closest to the origin, the switching curve that we have characterized above. Therefore, we have $\sigma(0) = 0$ and C_1' and C_2' as given in (5.6.5). If instead of letting the initial point (m_{10}, m_{20}) be on the switching curve (SC) that is close to the origin, we had assumed it is on the SC far from the origin, then there would exist $t' \in (0, T]$ for which $\sigma(0) = \sigma(t') = 0$. Since the final time is subject to optimization we might consider the time interval $[-t', T - t']$, and arguing similarly, the initial point being on the first SC at time $-t'$ implies the second switch is found at $t = 0$. Therefore, in order for a second switch to exist, there must exist $t' > 0$ such that $\sigma(-t') = 0$. We will however show below that $\sigma(t)$ is strictly increasing for all $t < 0$, and hence a second switch never happens.

To prove that $\sigma(t)$ is strictly increasing for all $t < 0$, let us show that $\mathrm{d}\sigma(t)/\mathrm{d}t > 0$ for all $t < 0$. From (5.6.5) we have that $C_2' \leq 0$ for all $m_{20}, m_{10} > 0$, then

$$\theta_1 \mu_1 C_1' \mathrm{e}^{\theta_1 t} - \theta_2 \mu_2 C_2' \mathrm{e}^{\theta_2 t} > 0 \Leftrightarrow \frac{\theta_1 \mu_1 C_1'}{\theta_2 \mu_2 C_2'} < \mathrm{e}^{(\theta_2 - \theta_1) t}. \tag{5.6.8}$$

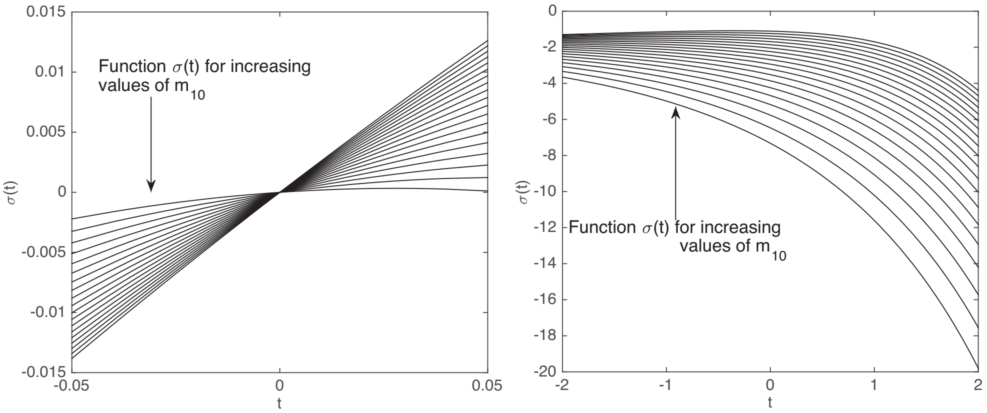


Figure 5.11: Left: Switching function $\sigma(t)$ for the case in which $\tilde{c}\mu/\theta$ and $\tilde{c}\mu$ have different ordering. Right: Switching function $\sigma(t)$ for the case in which $\tilde{c}\mu/\theta$ and $\tilde{c}\mu$ have same ordering.

Substituting the values of C_1' and C_2' after some algebra we obtain

$$\frac{\theta_1\mu_1 C_1'}{\theta_2\mu_2 C_2'} \leq 1 \Leftrightarrow m_{20} \geq (1-\rho_1-\rho_2)\frac{\mu_2}{\theta_1\theta_2}\left(\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\frac{\tilde{c}_2\mu_2}{\theta_2} - \frac{\tilde{c}_1\mu_1}{\theta_1}}\right) - \frac{\mu_2}{\mu_1}m_{10}.$$

It now suffices to plug the expression of m_{20} as given in (5.6.7) and prove the latter inequality. Note also that the right hand side of the inequality equals $\frac{a_1}{a_4} + \frac{\lambda_2}{\theta_2} + \frac{a_3}{a_4} - \frac{a_2\theta_2}{a_4\mu_1(1-\rho_1)} - \mu_2 m_{10}/\mu_1$ and $-\mu_2/\mu_1 = a_3\theta_2/(a_4(\mu_1-\lambda_1))$. Then after some algebra the inequality reduces to

$$\begin{aligned}&\frac{\frac{\mu_1}{\theta_1}\frac{\tilde{c}_2\mu_2}{\theta_2}(1-\rho_1-\rho_2)}{\frac{\tilde{c}_1\mu_1}{\theta_1} - \frac{\tilde{c}_2\mu_2}{\theta_2}}\left(1 + \frac{\theta_2 m_{10}}{\mu_1-\lambda_1} - \left(\frac{\theta_1 m_{10} + \mu_1 - \lambda_1}{\mu_1-\lambda_1}\right)^{\frac{\theta_2}{\theta_1}}\right)\\ &\leq m_{10}\left(1 + \frac{\theta_2 m_{10}}{\mu_1-\lambda_1} - \left(\frac{\theta_1 m_{10} + \mu_1 - \lambda_1}{\mu_1-\lambda_1}\right)^{\frac{\theta_2}{\theta_1}}\right),\end{aligned} \tag{5.6.9}$$

since $\tilde{c}_2\mu_2/\theta_2 > \tilde{c}_1\mu_1/\theta_1$ and $m_{10} \geq 0$. Define

$$\ell(m_{10}) := \left(1 + \frac{\theta_2 m_{10}}{\mu_1-\lambda_1} - \left(\frac{\theta_1 m_{10} + \mu_1 - \lambda_1}{\mu_1-\lambda_1}\right)^{\frac{\theta_2}{\theta_1}}\right).$$

Then for Inequality (5.6.9) to hold it suffices to show that $\ell(m_{10}) \geq 0$ for all $m_{10} \geq 0$. Observe that $\ell(0) = 0$ and $\ell(\cdot)$ is non-decreasing for all $m_{10} > 0$, therefore $\ell(m_{10}) \geq 0$ for all $m_{10} \geq 0$. This proves that $\theta_1\mu_1 C_1'/(\theta_2\mu_2 C_2') \leq 1$ and hence $\sigma(t)$ strictly increasing for all $t < 0$ which implies no other switch happens.

Now, since we have assumed $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ and $\tilde{c}_2\mu_2/\theta_2 > \tilde{c}_1\mu_1/\theta_1$ we have $\theta_1 > \theta_2$. The latter implies $\mathrm{e}^{(\theta_2-\theta_1)t} > 1$ for all $t < 0$, hence (5.6.8) is satisfied for all $t < 0$, see Figure 5.11, where we observe that $\sigma(t)$ has one root when $\tilde{c}\mu/\theta$ and $\tilde{c}\mu$ have opposite ordering.

Part III

Optimal dynamic control of stochastic systems

Chapter 6

Batch queues with abandonments

Contents

6.1	**Introduction**	**131**
6.2	**Model description**	**132**
6.3	**Optimality of threshold policies**	**135**
6.4	**Steady-state distribution**	**135**
6.5	**Characterization of the optimal threshold**	**138**
6.6	**Examples**	**139**
6.7	**Appendix**	**142**

In this chapter we consider a content delivery problem in which customers are processed in batches and may abandon before their service has been initiated. We model the problem as a Markovian single-server queue and analyze two different settings: (1) the system is cleared as soon as the server is activated, *i.e.*, the service rate is infinite, and (2) the service rate is finite. The objective is to determine the optimal clearing strategy that minimizes the average cost incurred by holding customers in the queue, having customers renege, and performing set-ups. This last cost is incurred upon activation of the server in the case of infinite service rate, and per unit of time the server is active otherwise. From the perspective of abandonments it is wise to serve customers in small batches; but it is profitable to accumulate customers since service comes at a very high cost. Our goal is to find the optimal balance between this service and abandonment cost trade-off.

In previous chapters we have concentrated on obtaining solutions that approximate the optimal solution, due to the complexity of the models under study. In this chapter we focus on a single class of customers, which enables to characterize an optimal solution explicitly.

6.1 Introduction

In this chapter we investigate a system that combines batch services with abandonment of customers. This model consists of an M/M/1 queue with an adapted service process in which customers may be delayed for batching of service. We consider two different settings: (1) the system is cleared immediately when a batch is taken into service, and (2) the service time is exponentially distributed with positive mean $1/\mu$. The service time of a batch is independent of the number of customers in the batch (multi-cast). Delays due to batching come at the cost of abandonment. In particular, customers may abandon the system while waiting to be served (expiration of their deadlines), for which we penalize the system at a fixed cost per abandoning customer. Such penalties can either represent the loss of the customer or the cost of serving the customer on an expensive back-up service. The abandonment process is modeled assuming exponential expiration times for individual customers. A similar methodology was adopted in Jean-Marie *et al.* [59] to investigate a system in which customers are batched for service to avoid a service set-up cost, but still must be served individually.

Scheduling multi-cast traffic with deadlines has various applications, *e.g.*, wireless sensor networks and video streams over cellular networks. Considerable attention has been given in the literature to systems where the specific deadlines of requests are known when the requests are made. This gives rise to well studied scheduling problems for queues and networks of queues with deadline-aware scheduling disciplines like Earliest Deadline First (EDF). For example, earliest deadline first queues are investigated under heavy traffic conditions in Doytchinov *et al.* [42] and Kruk *et al.* [65]. The optimality of EDF in terms of numbers of customers that meet their deadline was shown in Towsley *et al.* [89], assuming exponentially distributed service requirements. EDF and related schedulers assume that there is a separate "service" for each customer, whereas in our setting similar requests can be bundled.

In our first contribution we explore the optimality of threshold policies, which in this setting reduce to threshold-based policies. Structural results were obtained in Papadaki *et al.* [75] for a batch queue with finite service capacity and no abandonments. The presence of abandonments in the queue causes the system to be non uniformizable, and hence proving optimality of threshold policies requires methods like the SRT proposed in Bhulai *et al.* [25], we refer to Section 1.3.3 for a brief discussion on uniformization techniques. In our second contribution we analyze the stationary behavior of the system and compute the steady-state queue-length distribution for all $\mu \in \mathbb{R}_+ \cup \{\infty\}$, using a generating function approach. The latter allows us to find the minimum set-up cost, for each possible state of the system, such that taking a batch to service or staying idle is equally appealing cost-wise. For the infinite service speed case, we have been able to completely characterize the optimal threshold. In the finite service speed case, the analysis is more involved and a complete characterization could not be proven; instead the optimal threshold is determined by a function that is subject to optimization.

The remainder of the chapter is structured as follows. In Section 6.2 we describe the model. In Section 6.3 we prove that policies of threshold type are an optimal solution. In Section 6.4 we compute the steady-state queue-length distribution, which enables us in Section 6.5 to characterize the optimal threshold policy. Section 6.6 illustrates the obtained solution and its features through different numerical examples. Most of the proofs can be found in Appendix 6.7.

6.2 Model description

We consider an M/M/1 queue with batch service, infinite service capacity and customers abandonment. customers arrive to the queue according to a Poisson process with rate λ and have an exponentially distributed service requirement with mean $1/\mu$, which is independent of the batch size. customers that

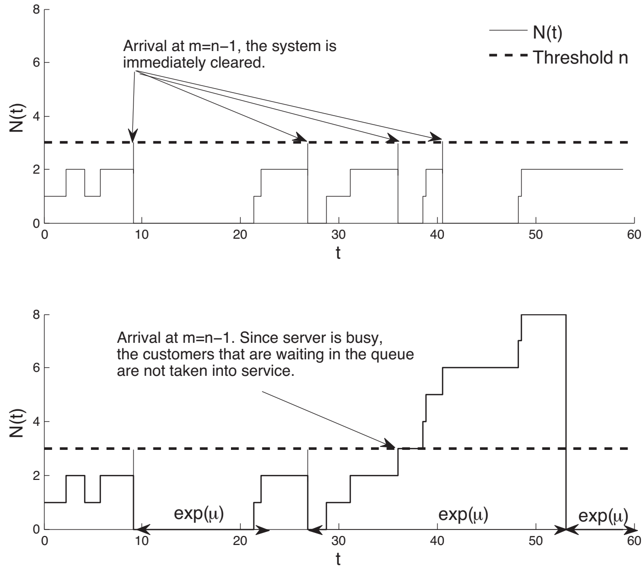


Figure 6.1: Simulation of process $N(t)$, number of waiting customers at time t, under threshold $n = 3$. Above the $\mu = \infty$ case, below the $\mu < \infty$ case. Below $exp(\mu)$ refers to the busy period of the server, which is exponentially distributed with rate $\mu < \infty$. As a consequence $N(t)$ not only depends on n but also on the length of each busy period.

are waiting in the queue abandon after an exponentially distributed amount of time with mean $1/\theta$. Furthermore, all interarrival times, service requirements and abandonment times are independent.

In every decision epoch the policy ϕ chooses whether to process the customers waiting in the queue or not. Once a customer has been admitted for service we assume that it can not abandon the system. Let $N^\phi(t) \in \{0, 1, \ldots\}$ denote the number of customers waiting in the queue at time t under the policy ϕ. Let $S^\phi(N^\phi(t)) \in \{0, 1\}$ denote the decision at time t under policy ϕ when there are $N^\phi(t)$ customers present in the queue. Namely, $S^\phi(N^\phi(t)) = 0$ if the server idles, and $S^\phi(N^\phi(t)) = 1$ if the server decides to take a batch into service. Due to the infinite capacity of the server we assume that, as soon as the server is activated, *i.e.*, $S^\phi(N^\phi)(t) = 1$, all customers that are waiting in the queue initialize their service. Hence, the batch size upon activation equals the number of customers waiting in the queue, $N^\phi(t)$.

We will analyze this problem in two different settings (see Figure 6.1, where ϕ is considered to be a threshold policy):

- The system is cleared as soon as the decision of taking a batch into service is made, that is, $\mu = \infty$.
- The service requirement are exponentially distributed with rate $\mu < \infty$.

In the first case ($\mu = \infty$), since service is immediate, the server empties the system as soon as the decision of activating the service is made. This means that whenever the policy ϕ decides to activate service, a batch of size $N^\phi(t)$ (*i.e.*, all customers waiting in the queue) will be instantaneously processed. In the second case ($\mu < \infty$), upon activation the server takes a batch of size $N^\phi(t)$ into service, and allocates an exponentially distributed amount of time to process it. While the server is busy, new customers might arrive to the queue. In this case, the server is not allowed to take a new batch into service until service completion of the previous batch; see Figure 6.1 (below) around $t = 37$. Hence, the evolution of $N^\phi(t)$ depends on both the policy ϕ and the state of the server (*i.e.*, idling or busy). In the $\mu = \infty$ case, the state of the system reduces to the number of customers waiting in the queue. In the $\mu < \infty$ case, the state of the system is given by (m, a), where m denotes the number of customers waiting in the queue and $a \in \{0, 1\}$ indicates whether the server is busy ($a = 1$) or idles ($a = 0$).

Let us denote by c the cost per unit of time customers are held in the queue, δ the penalty for customers abandoning the queue, by c_s^∞ the cost for setting up a service in the $\mu = \infty$ case and c_s the cost per unit of time the server is busy in the $\mu < \infty$ case. The objective of the present work is to find the policy ϕ so as to minimize

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\int_0^T \big[cN^\phi(t)\big]\,\mathrm{d}t + c_s^\infty G^\phi(T) + \delta R^\phi(T)\Big],$$

in the case $\mu = \infty$ and

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\int_0^T \big[cN^\phi(t) + c_s S^\phi(N^\phi(t))\big]\,\mathrm{d}t + \delta R^\phi(T)\Big],$$

in the $\mu < \infty$ case. In the latter objective functions $R^\phi(T)$ denotes the number of customers that abandoned in the interval $[0, T]$, and $G^\phi(T)$ the number of times the service has been activated in the interval $[0, T]$. Define $\mathcal{M}^\phi = \{m \in \{0, 1, \ldots\} : S^\phi(m) = 0, S^\phi(m+1) = 1\}$. By Dynkin's formula Anderson [3, Chapter 6.5], we have $\mathbb{E}[R^\phi(T)] = \theta\mathbb{E}[\int_0^T N^\phi(t)\mathrm{d}t]$, $\mathbb{E}[G^\phi(T)] = \lambda\mathbb{E}[\int_0^T \mathbf{1}_{\{N^\phi(t)\in\mathcal{M}^\phi\}}dt]$, and therefore the latter objective function is equivalent to finding ϕ that minimizes

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T [\tilde{c}N^\phi(t) + c_s^\infty \lambda \mathbf{1}_{\{N^\phi(t)\in\mathcal{M}^\phi\}}]\mathrm{d}t\right],$$

in the $\mu = \infty$ case, and

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T [\tilde{c}N^\phi(t) + c_s S^\phi(N^\phi(t))]\mathrm{d}t\right],$$

if $\mu < \infty$, where $\tilde{c} = c + \delta\theta$. Due to ergodicity of the system, the time-average optimal policy is equivalent to the optimal policy in steady-state, and hence we want to find ϕ such that

$$\min_\phi \left(\tilde{c}\mathbb{E}[N^\phi] + c_s^\infty \lambda \mathbb{E}(\mathbf{1}_{\{N^\phi\in\mathcal{M}^\phi\}})\right),$$

when $\mu = \infty$, and

$$\min_\phi \left(\tilde{c}\mathbb{E}[N^\phi] + c_s \mathbb{E}(\mathbf{1}_{\{S^\phi(N^\phi)=1\}})\right),$$

when $\mu < \infty$. The problem described above is a Markov Decision Process and we refer to it as Problem P throughout the chapter.

In general these types of problems are very complicated to solve due to the curse of dimensionality and the unbounded transition rates which yield a non-uniformizable system. However, we will observe in Section 6.3 that the result in Bhulai *et al.* [25] allows structural results to be proven for Problem P. We will show that in both settings, $\mu = \infty$ and $\mu < \infty$, threshold policies are average optimal. That is, there exists a threshold n for which the system prescribes not to take customers into service for all states $m \le n - 1$, and serves them otherwise.

6.3 Optimality of threshold policies

Obtaining structural results for non-uniformizable problems is rather involved. We have seen in Chapter 3 that under certain conditions one can prove optimality of threshold policies using the SRT approach proposed in Bhulai *et al.* [25], see also Section 1.3.3 for a brief discussion. In this section we are going to prove optimality of threshold polices using this technique. To do so we will prove that the requirements to apply the SRT, are satisfied herein.

Let us first define the finite state MDP, that is, let $L < \infty$ and let the number of customers waiting in the queue be $m \in \{0, \ldots, L\}$. Furthermore, let us smooth the arrival rate as follows: $q^{\phi,L}(m, m+1) = \lambda\left(1 - \frac{m}{L}\right),$ for all $L \geq m \geq 0$. This truncation now yields bounded transition rates, and the smoothing of the arrival rate guarantees the structure of the original value function to be maintained.

Having defined the finite state space MDP, in Appendix 6.7.1, we show that the conditions in Bhulai *et al.* [25, Theorem 3.1] hold for Problem P, making SRT suitable for the model under study.

We next prove optimality of threshold policies for $\mu \in \mathbb{R}^+ \cup \{\infty\}$. The proof for the case $\mu = \infty$ follows from Proposition 2.1, and the proof for $\mu < \infty$ can be found in Appendix 6.7.1 together with the verification of the conditions in Bhulai *et al.* [25, Theorem 3.1].

Proposition 6.1. *Letting $\mu \in \mathbb{R}^+ \cup \{\infty\}$, then there exists a threshold policy n that is optimal for Problem P.*

6.4 Steady-state distribution

In this section we derive the steady-state distribution of Problem P under policies of threshold type. We will refer to these policies as threshold n or simply $\phi = n$, when the threshold is determined by $n \in \{0, 1, \ldots\}$. Recall that threshold n prescribes to be passive for all states $m \leq n-1$, and active for all $m \geq n$. Throughout this section, for clarity of exposition, we drop the dependency on n from the notation. The transition rates for this Markov process are as follows:

$$\begin{aligned}
&q(m, m+1) = \lambda, \text{ for all } 0 \leq m \leq n-1,\\
&q(m, m-1) = \theta m, \text{ for all } 0 < m \leq n-1,\\
&q(n, 0) = \mu,
\end{aligned}$$

in the $\mu = \infty$ case, and

$$\begin{aligned}
&q((m,0),(m+1,0)) = \lambda, \text{ for all } 0 \leq m \leq n-1, q((m,1),(m+1,1)) = \lambda, \text{ for all } 0 \leq m,\\
&q((m,0),(m-1,0)) = \theta m, \text{ for all } 0 < m \leq n-1, q((m,1),(m-1,1)) = \theta m, \text{ for all } 0 < m,\\
&q((n-1,0),(0,1)) = \lambda,\\
&q((m,1),(0,1)) = \mu, \text{ for all } m \geq n, q((m,1),(m,0)) = \mu, \text{ for all } 0 \leq m \leq n-1,
\end{aligned}$$

for the $\mu < \infty$ case.

In Section 6.4.1 we derive the steady-state distribution of Problem P for the case $\mu = \infty$ and in Section 6.4.2 for the case $\mu < \infty$.

6.4.1 Infinite service rate

In the case of infinite service speed we assume that the system empties the queue as soon as the decision to serve is made. In this setting, under threshold n, the state space is given by $E = \{0, 1, \ldots, n-1\}$. As soon as an arrival occurs in state $m = n-1$ the system is immediately cleared.

Let us define π_m as the steady-state probability of being in state $m \in \{0, \ldots, n-1\}$. Hence, we have the following balance equations

$$\lambda \pi_{m-1} = \theta m \pi_m + \lambda \pi_{n-1}, \quad \forall\ 0 < m \leq n-1,$$

together with the normalizing equation $\sum_{i=0}^{n-1} \pi_i = 1$. We now solve the balance equations, that is,

$$\begin{aligned}
\pi_m &= (m+1)\frac{\theta}{\lambda}\pi_{m+1} + \pi_{n-1} \\
&= \pi_{n-1} + (m+1)\frac{\theta}{\lambda}\left((m+2)\frac{\theta}{\lambda}\pi_{m+2} + \pi_{n-1}\right) \\
&= \pi_{n-1}\left(1 + \frac{\theta}{\lambda}(m+1)\right) + \pi_{m+2}\left(\frac{\theta}{\lambda}\right)^2 (m+1)(m+2) \\
&= \pi_{n-1}\left(1 + \frac{\theta}{\lambda}(m+1) + \left(\frac{\theta}{\lambda}\right)^2 (m+1)(m+2)\right) \\
&\quad + \pi_{m+3}\left(\frac{\theta}{\lambda}\right)^3 (m+1)(m+2)(m+3) \\
&= \ldots \\
&= \pi_{n-1}\left(1 + \sum_{i=1}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i \frac{(m+i)!}{m!}\right).
\end{aligned}$$

Hence,

$$\pi_m = \pi_{n-1}\left(\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i \frac{(m+i)!}{m!}\right),$$

for all $m = 0, 1, \ldots, n-1$. Moreover, from the normalization equation, we obtain

$$\pi_{n-1} = \left(\sum_{m=0}^{n-1}\left(\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i \frac{(i+m)!}{m!}\right)\right)^{-1}. \tag{6.4.1}$$

We therefore have obtained the expression for all π_m and $0 \leq m \leq n-1$.

6.4.2 Finite service rate

In the case of finite service speed, the state of the system is given by (m, a) where m denotes the number of customers waiting in the queue and $a \in \{0, 1\}$ whether the server is available ($a = 0$) or busy ($a = 1$). Observe that under threshold n, during the idle period of the server, the number of customers in the queue, *i.e.*, m, takes values in the set $\{0, \ldots, n-1\}$. The latter means that, as soon as an arrival happens in state $m = n-1$, the server activates the service with a batch of n customers. Once the server is active, during the time the batch is being processed (exponentially distributed with mean $1/\mu$), the number of customers waiting in the queue are such that $m \in \mathbb{N} \cup \{0\}$. That is, even if the threshold n is reached, the activation of the server is postponed until it completes processing the previous batch.

Let us denote by $\pi(m,a)$ the steady-state probability of being in state (m,a), for all $m \geq 0$ and $a \in \{0,1\}$, and assume $\pi(m,0) = 0$ for all $m \geq n$. As in the previous section, we omit the dependence on n from the notation. Then, $\pi(m,a)$ for all $m \in \mathbb{N} \cup \{0\}$ and $a \in \{0,1\}$ can be derived from the following balance equations: for all $m \in \mathbb{N}$

$$(\lambda + m\theta + \mu)\pi(m,1) = \lambda\pi(m-1,1) + (m+1)\theta\pi(m+1,1), \tag{6.4.2}$$

and for all $0 \leq m \leq n-1$

$$(\lambda + m\theta)\pi(m,0) = \lambda\pi(m-1,0) + \mu\pi(m,1) + (m+1)\theta\pi(m+1,0). \tag{6.4.3}$$

In order to solve the balance equations in (6.4.2) and (6.4.3) we will use their corresponding ordinary generating functions. Observe in Equation (6.4.3) that the steady-state probabilities of the idle period depend on the steady-state probabilities of the busy period. Therefore, we first obtain the closed-form expression of $\pi(m,1)$ for all m, and using these expressions we derive those that correspond to the idle period, *i.e.*, $\pi(m,0)$ for all $n-1 \geq m \geq 0$. The explicit expression of the probabilities are presented in Propositions 6.2. The calculations to derive these expressions can be found in Appendix 6.7.2, for the busy period, and in Appendix 6.7.2, for the idle period.

Proposition 6.2. *Let* $a_1(0) := 1$,

$$a_1(1) := \frac{\lambda+\mu}{\theta} - \frac{\mathrm{e}^{\lambda/\theta}}{\sum_{i=0}^{\infty} \frac{(\lambda/\theta)^i}{i!(\mu/\theta+i)}}, \text{ and,}$$

$$a_1(m) := \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}\left(\frac{\sum_{j=0}^{\infty}\frac{\left(\frac{\lambda}{\theta}\right)^j(-\ell_{k-1}(-\frac{\mu}{\theta}-j+1))}{j!}}{\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)}}\sum_{i=0}^{m-k}\binom{m-k}{i}\left(\frac{\lambda}{\theta}\right)^{m-k-i}\ell_i\left(\frac{\mu}{\theta}\right)\right),$$

for all $m \geq 2$, *where* $\ell_k(x)$ *is the Pochhammer function or the rising factorial. Let* $a_0^n(n-1) := \frac{\mu}{\lambda}\sum_{m=0}^{n-1} a_1(m)$,

$$a_0^n(m) := \left(\frac{\mu}{\lambda}\sum_{r=0}^{n-1}a_1(r)\right)\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!} - \frac{\mu}{\lambda}\sum_{r=m+1}^{n-1}a_1(r)\left(\sum_{i=0}^{r-m-1}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\right),$$

for all $n-2 \geq m > 0$, *and* $a_0^n(0) := \frac{\mu}{\lambda} + \frac{\theta}{\lambda}a_0^n(1)$. *Then taking*

$$\pi(0,1) = \left(\sum_{m=0}^{n-1}a_0^n(m) + \sum_{m=0}^{\infty}a_1(m)\right)^{-1},$$

$\pi(m,0) = a_0^n(m)\pi(0,1)$, *and* $\pi(m,1) = a_1(m)\pi(0,1)$, *we obtain* $\sum_{m=0}^{n-1}\pi(m,0) + \sum_{m=0}^{\infty}\pi(m,1) = 1$, *and* $\pi(m,0)$ *and* $\pi(m,1)$ *solve Equations* (6.4.2) *and* (6.4.3).

Once we have obtained all $\pi(m, 0)$ for all $n-1 \geq m \geq 0$, and $\pi(m, 1)$ for all $m \geq 0$ we can proceed to compute the mean number of customers in the system under threshold policy n, as well as the mean amount of time the set-up cost is incurred. In the next section we show that this provides a characterization of the optimal threshold.

6.5 Characterization of the optimal threshold

In this section we characterize the optimal threshold policy using the steady-state probabilities that we have computed above. This characterization, as we will see, depends on the set-up cost c_s^∞ and service cost c_s. For the infinite service speed case ($\mu = \infty$), the characterization of the optimal threshold is explicit whose solution we present in Proposition 6.3. In the finite service speed case, the characterization of the optimal threshold is determined by a function that is subject to optimization; see Proposition 6.4.

From Proposition 6.1 we know that threshold policies are optimal for Problem P. Then Problem P can be rewritten as

$$\min_{n \in \{0,1,\dots\}} \left(\tilde{c}\mathbb{E}(N^n) + c_s^\infty \lambda \mathbb{E}(\mathbf{1}_{\{N^n = n-1\}})\right),$$

in the $\mu = \infty$ case, where $\mathbb{E}(\mathbf{1}_{\{N^n=n-1\}}) = \pi_{n-1}^n$ and

$$\min_{n \in \{0,1,\dots\}} \left(\tilde{c}\mathbb{E}(N^n) + c_s \mathbb{E}(\mathbf{1}_{\{S^n(N^n)=1\}})\right),$$

in the $\mu < \infty$ case, where $\mathbb{E}(\mathbf{1}_{\{S^n(N^n)=1\}}) = \sum_{m=0}^\infty \pi^n(m, 1)$. Let us introduce the following notation, we denote by $P_b^n = \pi_{n-1}^n$ in the $\mu = \infty$ case and $P_b^n = \sum_{m=0}^\infty \pi^n(m, 1)$ in the $\mu < \infty$ case. From now on we shall denote by $c_s = c_s^\infty \lambda$ in the $\mu = \infty$ case.

In the following proposition we propose an explicit representation of the optimal threshold. The proof can be found in Appendix 6.7.3.

Proposition 6.3. *Let us define $\alpha(n)$ such that*

$$\alpha(n) := \tilde{c}\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-1})}{P_b^{n-1} - P_b^n}.$$

for $n > 1$. If $\alpha(n)$ is non-decreasing in n, P_b^n is non-increasing in n, and if $\alpha(n) \leq c_s < \alpha(n+1)$, then n is optimal for Problem P.

In the following lemma we prove that the assumptions in Proposition 6.3 hold in the case $\mu = \infty$. The proof can be found in Appendix 6.7.4.

Lemma 6.1. *Let $\mu = \infty$, and π_m^n as given in Section 6.4.1 after adding the superscript n. Then*

- *$P_b^n = \pi_{n-1}^n$ is convex non-increasing.*
- *The function $\alpha(n)$, as defined in Proposition 6.3, is non-decreasing.*

Corollary 6.1. *Assume $\mu = \infty$, $c_s = c_s^\infty \lambda$ and define $\alpha(1) := -\infty$. Then, if $\alpha(n) \leq c_s < \alpha(n+1)$ for some $n \geq 1$, n is the optimal threshold policy for Problem P.*

Proof. The proof follows from Proposition 6.3 and Lemma 6.1. □

In the case $\mu < \infty$ we could not prove α to be non-decreasing. The optimal threshold then has to be characterized differently. This characterization is given in the following proposition.

Proposition 6.4. *Let $\mathcal{N}_i = \mathbb{N} \cup \{0\} \backslash \{0, \ldots, n_i\}$ for a given n_i, let P_b^n be non-increasing, and define $\beta(\cdot)$ as follows: Define $n_0 = 0$ and*

Step i. Compute

$$\beta_i := \inf_{n \in \mathcal{N}_{i-1}} \tilde{c} \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n_{i-1}})}{P_b^{n_{i-1}} - P_b^n}, i \geq 1, \tag{6.5.1}$$

and denote by n_i the largest $n \in \mathcal{N}_{i-1}$ such that (6.5.1) is minimized and define $\beta(n) = \beta_i$ for all $n_i > n \geq n_{i-1}$. If $n_i = \infty$ stop and let $\beta(n) = \beta_i$ for all $n \geq n_i$, otherwise jump to step $i+1$.

Then, β_i is strictly non-decreasing in i and if $\beta_i \leq c_s < \beta_{i+1}$, then n_i is optimal for Problem P. Moreover, if $c_s < \beta_1$, then it will be optimal to always serve.

Proof. Let us first prove that β_i is non-decreasing in i. Recall that by definition

$$\frac{\mathbb{E}(N^{n_i}) - \mathbb{E}(N^{n_{i-1}})}{P_b^{n_{i-1}} - P_b^{n_i}} < \frac{\mathbb{E}(N^{n_{i+1}}) - \mathbb{E}(N^{n_{i-1}})}{P_b^{n_{i-1}} - P_b^{n_{i+1}}},$$

and then $(\mathbb{E}(N^{n_i}) - \mathbb{E}(N^{n_{i-1}}))(P_b^{n_{i-1}} - P_b^{n_{i+1}}) < (\mathbb{E}(N^{n_{i+1}}) - \mathbb{E}(N^{n_{i-1}})))(P_b^{n_{i-1}} - P_b^{n_i})$. Upon adding and subtracting $\mathbb{E}(N^{n_i})(P_b^{n_{i-1}} - P_b^{n_i})$ on the RHS, and after some algebra, we obtain $\beta_i < \beta_{i+1}$.

Having proven that β_i is non-decreasing, the optimality of threshold n_i if $\beta_i \leq c_s < \beta_{i+1}$ can be proven in the same way as in Proposition 6.3. □

The following conjecture establishes the characterization proposed in Proposition 6.4 to hold when $\mu < \infty$. The justification can be found in Appendix 6.7.5.

Conjecture 6.1. *Assume $\mu < \infty$ and $\pi^n(m, a)$ as given in Section 6.4.2, for all $m \geq 0$ and $a \in \{0, 1\}$, after adding the superscript n. Then, $P_b^n = \pi^n(0, 1) \sum_{m=0}^{\infty} a_1(m)$, is non-increasing.*

We characterize the optimal solution in the case $n_i = i$ in Proposition 6.4 in the following conjecture for the case $\mu < \infty$.

Conjecture 6.2. *Assume $\mu < \infty$ and assume*

$$\tilde{c} \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-1})}{P_b^{n-1} - P_b^n}, \tag{6.5.2}$$

to be non-decreasing. Then $\beta(n)$ is defined by (6.5.2) and is non-decreasing. Hence if $\beta(n) \leq c_s < \beta(n+1)$ then threshold n is optimal. If $c_s < \beta(1)$, then 0 is the optimal threshold (always serve).

Proof. The proof follows from Proposition 6.4 and Conjecture 6.1. □

In the next section we analyze the optimal threshold policies in both frameworks.

6.6 Examples

In this section we illustrate the features of the optimal threshold policies that have been characterized in Section 6.5 through different examples. In Examples 1 and 2 we illustrate the optimal threshold policy for

Table 6.1: Example 1: Minimum set-up cost c_s^∞ such that n is optimal.

n	1	2	3	4	5	6
c_s^∞	$-\infty$	0.6096	2.4359	6.2595	12.8192	22.5343

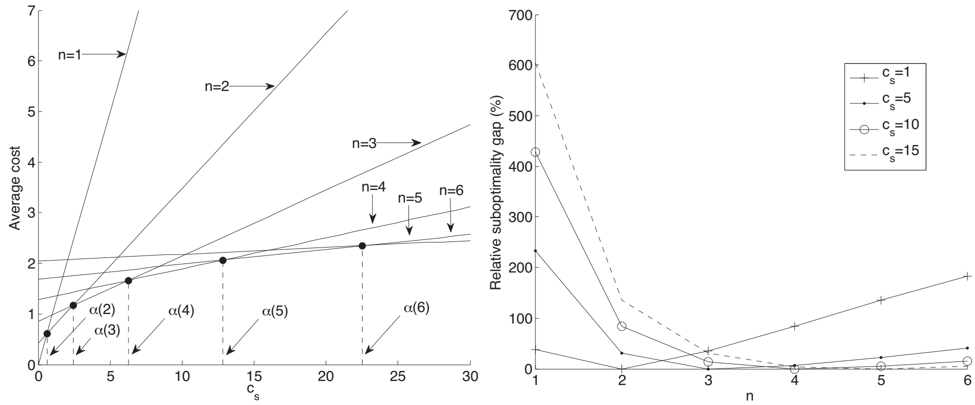


Figure 6.2: On the left the average cost under different threshold policies and varying value of c_s. On the right the relative suboptimality gap of different threshold policies with respect to the optimal threshold.

different values of the set-up cost c_s^∞ and service cost c_s and we evaluate the performance of non-optimal threshold policies in comparison with the optimal one for the case $\mu = \infty$ and $\mu < \infty$, respectively. In Examples 3 and 4 we consider the influence on the optimal threshold policy for varying values of θ in the case $\mu = \infty$ and for varying values of θ and μ in the case $\mu < \infty$.

Example 1: Let us assume $\lambda = 4$, $\mu = \infty$, $\theta = 1.5$ and $\tilde{c} = 1$. Then the minimum value of c_s, where $c_s = c_s^\infty \lambda$, such that n is optimal for Problem P, is given by $\alpha(n)$, whose values are presented in Table 6.1. In Figure 6.2 (left) we illustrate this optimal solution, where we plot the average cost $\mathbb{E}(N^n) + c_s P_b^n$ incurred by policy n, for different values of c_s. We only present the solution up to $n = 6$, noting that a characterization for all n can be found. In Figure 6.2 (right) we present the relative sub optimality gap of non-optimal threshold policies with respect to the optimal that we have just characterized. Observe that their performance is very poor.

Example 2: Let us assume $\lambda = 2$, $\mu = 0.5$, $\theta = 0.5$ and $\tilde{c} = 1$. Then the minimum value of c_s, such that n is optimal for Problem P is given by $\beta(n)$ which in this case is given by (6.5.2). The values of $\beta(n)$ are presented in Table 6.2 for n up to 5. We observe that under the assumption $c_s > 0$ thresholds $n = 0, 1, 2$ are never optimal in this particular example, which means that when there are 1 or 2 customers waiting to be served the server will idle. These results are illustrated in Figure 6.3 (left), where we plot the average cost $\mathbb{E}(N^n) + c_s P_b^n$ incurred by policy n for different values of c_s. In Figure 6.3 (right) we plot the relative sub optimality gap of different threshold policies with respect to the optimal threshold, and observe that non-optimal thresholds incur a huge cost.

Table 6.2: Example 2: Minimum set-up cost c_s such that n is optimal.

n	1	2	3	4	5
c_s	-2	-1.151	-0.2581	0.7157	1.7937

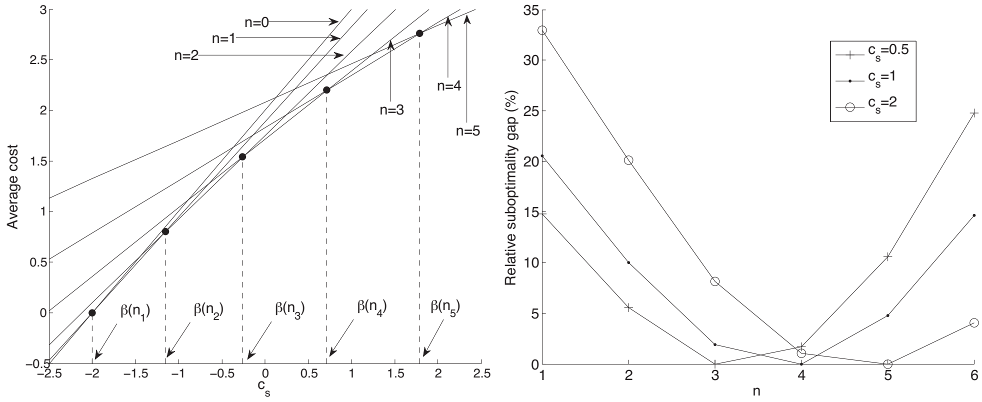


Figure 6.3: On the left the average cost under different threshold policies and varying value of c_s. On the right the relative suboptimality gap of different threshold policies with respect to the optimal threshold.

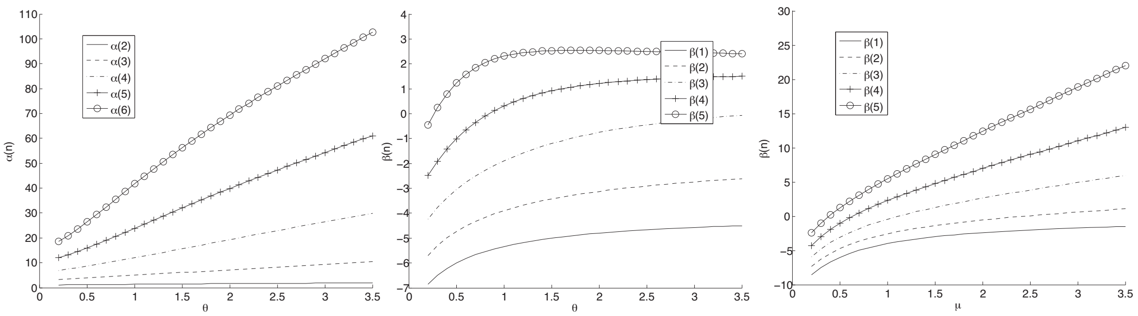


Figure 6.4: We plot the values of α and β for varying parameters θ and μ. On the left the case $\mu = \infty$ for varying values of θ. In the middle and on the right the case $\mu < \infty$ for varying values of θ and μ respectively.

In the following two examples we observe the policy changes as θ and μ vary.

Example 3: Let us assume $\lambda = 4$, $\mu = \infty$ and $c = \delta = 1$ and let θ vary in the interval $[0, 3.5]$. We observe in Figure 6.4 (left) that the minimum value of c_s, where $c_s = c_s^\infty \lambda$, such that threshold policy n is optimal increases as θ increases. This means that for a fixed set-up cost and for increasing abandonment rates, the system prescribes to activate service earlier to avoid incurring the abandonment penalty.

Example 4: Let us assume $\lambda = 2$ and $c = \delta = 1$. We first consider the case $\mu = 0.5$ and let θ vary in the interval $[0, 3.5]$; see Figure 6.4 (middle). We observe that as θ grows large, $\beta(n)$ becomes constant. This phenomenon is explained by the fact that abandonments happen both in the busy period and in the idle period of the server, and hence for a fixed set-up cost the threshold is maintained as θ grows. Secondly, we consider $\theta = 0.5$ and let μ vary in the interval $[0, 3.5]$. We observe in Figure 6.4 (right) that the faster the service, the smaller the optimal threshold for a fixed c_s.

6.7 Appendix

6.7.1 Proof of Proposition 6.1: the case $\mu < \infty$

We will first check that the conditions to apply SRT are satisfied by the model under study. Later on we will prove that threshold type of policies are optimal for $\mu < \infty$.

Conditions to be checked for Bhulai *et al.* [25, Th. 3.1]

We prove that the conditions to be checked for Bhulai *et al.* [25, Th. 3.1] are satisfied by Problem P. Let $E = \mathbb{N} \cup \{0\}$, and define $h(m) = \mathrm{e}^{\epsilon m}$, then, by Definition 3.1, h is a moment function. For the statement of the conditions that must be verified we refer to Section 3.7.1, where we have verified this conditions for the multi-class abandonment queue.

Let us assume $\mu < \infty$ and $k_1 = (k_{11}, k_{12})$, then the first condition, as stated in Section 3.7.1, reduces to finding $\epsilon > 0$ and $\tilde{M} > 0$ such that for all $m \geq \tilde{M}$

$$\begin{aligned}
&\lambda\left(1-\frac{m}{L}\right)\mathrm{e}^{\epsilon(m+1)}(1-S^{\phi}(m)) + \theta m \mathrm{e}^{\epsilon(m-1)}(1-S^{\phi}(m)) + \left(\lambda\left(1-\frac{m}{L}\right)+\theta m\right)S^{\phi}(m)\\
&-\left(\lambda\left(1-\frac{m}{L}\right)+\theta m\right)\mathrm{e}^{\epsilon m} \leq -k_{11}\mathrm{e}^{\epsilon m},\\
&\mu \mathrm{e}^{\epsilon m} + \lambda\left(1-\frac{m}{L}\right)\mathrm{e}^{\epsilon(m+1)} + \theta m \mathrm{e}^{\epsilon(m-1)} - \left(\lambda\left(1-\frac{m}{L}\right)+\mu+\theta m\right)\mathrm{e}^{\epsilon m} \leq -k_{12}\mathrm{e}^{\epsilon m},
\end{aligned}$$

where k_{11} and k_{12} are constants. In the latter equation, the first inequality corresponds to the state $(m, 0)$ and the second inequality to $(m, 1)$. After some algebra these two inequalities reduce to

$$\lambda\left(1-\frac{m}{L}\right)(\mathrm{e}^{\epsilon}-1+S^{\phi}(m)(\mathrm{e}^{-\epsilon m}-\mathrm{e}^{\epsilon})) + \theta m(\mathrm{e}^{-\epsilon}-1+S^{\phi}(m)(\mathrm{e}^{-\epsilon m}-\mathrm{e}^{-\epsilon})) \leq -k_{11},$$

where $\lambda\left(1-\frac{m}{L}\right)\left(\mathrm{e}^{\epsilon}-1+S^{\phi}(m)(\mathrm{e}^{-\epsilon m}-\mathrm{e}^{\epsilon})\right)$ is upper bounded by a constant, say κ, and $\theta m(\mathrm{e}^{-\epsilon}-1+S^{\phi}(m)(\mathrm{e}^{-\epsilon m}-\mathrm{e}^{-\epsilon})) \leq 0$, hence, we can find $\tilde{M}$ large enough so that $\theta m(\mathrm{e}^{-\epsilon}-1+S^{\phi}(m)(\mathrm{e}^{-\epsilon m}-\mathrm{e}^{-\epsilon})) \leq -\kappa$. The second inequality writes

$$\lambda\left(1-\frac{m}{L}\right)(\mathrm{e}^{\epsilon}-1) + \theta m(\mathrm{e}^{-\epsilon}-1) \leq -k_{12},$$

where $\lambda\left(1-\frac{m}{L}\right)(\mathrm{e}^{\epsilon}-1)$ is upper bounded and can be made as negative as desired for a big enough $\tilde{M}$.

The second condition to be verified, as stated in Section 3.7.1, that is continuity of $q^{\phi,L}(\cdot,\cdot)$ in $S^{\phi}(N^{\phi}(t))$ and L, follows by construction.

Optimality of threshold policies

Let us denote by $V(m, a)$ the value function corresponding to Problem P in the case $\mu < \infty$, and let g be the average cost incurred by an optimal policy. Recall that the state of the system in this framework reduces to (m, a), where m is the number of customers waiting in the queue and $a \in \{0, 1\}$ denotes the state of the server, busy if $a = 1$ and available if $a = 0$. The value function $V(m, a)$ for all $m \geq 0$ and

$a \in \{0,1\}$ satisfies the Bellman equation, see Appendix A.1 (Equation A.1.1), namely

$$\begin{aligned}(\lambda+\mu+\theta m)V(m,0)+g =&\tilde{c}m+\min\{\lambda V(m+1,0)+\theta m V((m-1)^+,0)\\&+\mu V(m,0), c_s+\lambda V(1,1)+\theta m V(0,1)+\mu V(0,0)\},\end{aligned}$$

and

$$(\lambda+\mu+\theta m)V(m,1)+g = \tilde{c}m + c_s + \lambda V(m+1,1) + m\theta V((m-1)^+,1) + \mu V(m,0).$$

We will now prove that an optimal policy solving the Bellman equation is of threshold type, that is, if active action is optimal in m then active action is optimal in $m' \geq m$. In order to do so, let us first define

$$\begin{aligned}f(m,0) :=&\tilde{c}m + \lambda V(m+1,0) + \mu V(m,0) + \theta m V(m-1,0),\\ f(m,1) :=&\tilde{c}m + c_s + \lambda V(1,1) + \theta m V(0,1) + \mu V(0,0),\end{aligned}$$

and $\varphi(m) = \min\left(b \in \arg\min_{a\in\{0,1\}} f(m,a)\right)$. It then suffices to show that $\varphi(m') \geq \varphi(m)$ for $m' \geq m$. Let $a \geq \varphi(m')$. By definition of $\varphi(\cdot)$

$$f(m',\varphi(m')) - f(m',a) \leq 0. \tag{6.7.1}$$

Let us now prove that $-f(m,0)$ is submodular Puterman [77, Chapter 4.7.2], that is, for all $m' \geq m$ and $a \in \{0,1\}$

$$f(m',a) + f(m,\varphi(m')) \leq f(m',\varphi(m')) + f(m,a). \tag{6.7.2}$$

Assuming first the case $\varphi(m') = 0$ and $a = 0$, then (6.7.2) is trivially satisfied; similarly (6.7.2) is satisfied in the case $\varphi(m') = 1$ and $a = 1$. We are left with the case $\varphi(m') = 0$ and $a = 1$, for which (6.7.2) reduces to $f(m,0) - f(m,1) \leq f(m',0) - f(m',1)$. Since Problem P for $\mu < \infty$ is non-uniformizable, we use the SRT approach. We truncate the system with the parameter L, and define smoothed arrival transition rates, i.e., $q(m,m+1) = \lambda\left(1-\frac{m}{L}\right)$ for all $m \leq L$. We denote by $V^L(\cdot,\cdot)$ the value function of the truncated system. Since the conditions required to apply SRT are satisfied and Bhulai *et al.* [25, Theorem 3.1], we have that $V^L \to V$ as $L \to \infty$ and the value function maintains its structural properties. Assume, without loss of generality, $\lambda+\mu+\theta L = 1$. Then, applying the value iteration algorithm, see Section 1.3.3, to the truncated system, we have

$$\begin{aligned}V_{t+1}^L(m,0) = \tilde{c}m + \min\{&\lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0) + \left(\lambda\frac{m}{L}+\mu+\theta(L-m)\right)V_t^L(m,0) + \theta m V_t^L(m-1,0),\\ &c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0)\},\end{aligned} \tag{6.7.3}$$

and

$$\begin{aligned}V_{t+1}^L(m,1) &= \tilde{c}m + c_s + \lambda\left(1-\frac{m}{N}\right)V_t^L(m+1,1) + \mu V_t^L(m,0)\\ &+ \left(\lambda\frac{m}{N} + (N-m)\theta\right)V_t^L(m,1) + m\theta V_t^L(m-1,1),\end{aligned}$$

since $V_{t+1}^L(m,a) - V_t^L(m,a) = g$. It suffices to prove that $f_t^L(m,0) - f_t^L(m,1) \leq f_t^L(m',0) - f_t^L(m',1)$, where

$$
\begin{aligned}
f_t^L(m,0) =& \lambda \left(1 - \frac{m}{L}\right) V_t^L(m+1,0) + \theta m V_t^L(m-1,0) + \left(\lambda \frac{m}{L} + \mu + \theta(L-m)\right) V_t^L(m,0), \\
f_t^L(m,1) =& \tilde{c}m + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0).
\end{aligned}
$$

V^L is a non-decreasing function and $f_t^L(m,0) - f_t^L(m,1) \leq f_t^L(m',0) - f_t^L(m',1)$ for all $m' \geq m$ and all t, the proof of both of which can be found in below. By the value iteration argument $f_t^L \to f^L$ as $t \to \infty$, and from the SRT we have that $V^L \to V$ and $f^L \to f$ component-wise. Therefore, subadditivity (reverse inequality w.r.t. supermodularity) of of $V^L(\cdot,0)$ implies subadditivity of $V(\cdot,0)$. Having proven (6.7.1) and (6.7.2), and upon combining them, we have that for all $a \geq \varphi(m')$ and $m' \geq m$

$$
\begin{aligned}
f(m,\varphi(m')) &\leq f(m',\varphi(m')) - f(m',a) + f(m,a) \\
&\leq f(m,a).
\end{aligned}
$$

Hence $\varphi(m) \leq \varphi(m')$, which concludes the proof.

Proof for $V^L(\cdot,0)$ non-decreasing. To prove that $V^L(\cdot,0)$ is a non-decreasing function we define $V_0^L(m) = 0$ for all $m \leq L$. Given g the optimal average cost, and assuming w.l.o.g $\lambda + \mu + \theta L = 1$ the Bellman equation reads

$$
\begin{aligned}
& V_{t+1}^L(m,0) = \tilde{c}m + \min\Bigg(\lambda \left(1 - \frac{m}{L}\right) V_t^L(m+1,0) \\
& + \left(\lambda \frac{m}{L} + \mu + \theta(L-m)\right) V_t^L(m,0) + \theta m V_t^L(m-1,0); \\
& c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0)\Bigg),
\end{aligned}
\tag{6.7.4}
$$

and

$$
\begin{aligned}
V_{t+1}^L(m,1) =& \tilde{c}m + \lambda \left(1 - \frac{m}{L}\right) V_t^L(m+1,1) + \mu V_t^L(m,0) \\
& + \left(\lambda \frac{m}{L} + (L-m)\theta\right) V_t^L(m,1) + m\theta V_t^L(m-1,1),
\end{aligned}
\tag{6.7.5}
$$

since $V_{t+1}^L(m,0) - V_t^L(m,0) = g$.

We will first prove that for all t, $V_t^L(m,0)$ is non-decreasing in m. We will argue by induction: first we show that $V_0^L(m,0) \geq V_0^L(m',0)$ for all $m \geq m' \geq 0$ implies that $V_1^L(m,0) \geq V_1^L(m',0)$, and later on we will prove that assuming that $V_t^L(m,a)$ is non-decreasing in m, implies $V_{t+1}^L(m,1)$ to be non-decreasing. By definition $V^L(m,a)$ is the asymptotic difference in total reward from starting at state m instead of starting at the reference state which, without loss of generality, we set at 0. We choose $V_0^L(m,a) = 0$ for all $m \geq 0$ and $a \in \{0,1\}$, then, from (6.7.4) and (6.7.5) we obtain $V_1^L(m,a) = \tilde{c}m$. Since $\tilde{c} > 0$, $V_1^L(m,a) \geq V_1^L(m-1,a)$ for all $m \geq 1$. We now assume that $V_t^L(m,0)$ is non-decreasing, and we prove that $V_{t+1}^L(m,0) \geq V_{t+1}^L(m-1,0)$ for all $L \geq m \geq 1$, that is, after substitution of (6.7.4), equivalent to proving

$$\tilde{c}m+\min\biggl(\lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0)+\theta m V_t^L(m-1,0)+\left(\lambda\frac{m}{L}+\mu+\theta(L-m)\right)V_t^L(m,0),c_s+\lambda V_t^L(1,1)$$
$$+\theta N V_t^L(0,1)+\mu V_t^L(0,0)\biggr)$$
$$\geq\tilde{c}(m-1)+\min\biggl(\lambda\left(1-\frac{m-1}{L}\right)V_t^L(m,0)+\left(\lambda\frac{m-1}{L}+\mu+\theta(L-m+1)\right)V_t^L(m-1,0)$$
$$+\theta(m-1)V_t^L((m-2)^+,0),c_s+\lambda V_t^L(1,1)+\theta L V_t^L(0,1)+\mu V_t^L(0,0)\biggr). \tag{6.7.6}$$

We will now prove that Inequality (6.7.6) is satisfied for all possible action combinations in states m and $m-1$. Let us first assume that in both m and $m-1$ passive action is optimal, then (6.7.6) reduces to

$$\begin{aligned}&\tilde{c}m+\lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0)+\theta m V_t^L(m-1,0)+\left(\lambda\frac{m}{L}+\mu+\theta(L-m)\right)V_t^L(m,0)\\&\geq\tilde{c}(m-1)+\lambda\left(1-\frac{m-1}{L}\right)V_t^L(m,0)+\left(\lambda\frac{m-1}{L}+\mu+\theta(L-m+1)\right)V_t^L(m-1,0)\\&+\theta(m-1)V_t^L((m-2)^+,0),\end{aligned}$$

which after some calculations reads

$$\tilde{c}+\lambda\left(1-\frac{m}{L}\right)\Delta V_t^L(m+1,0)+\theta(m-1)\Delta V_t^L(m-1,0)+\left(\lambda\frac{m-1}{L}+\mu+\theta(L-m)\right)\Delta V_t^L(m,0)\geq 0,$$

for all $L\geq m\geq 1$ and $\Delta V_t^L(m,0)=V_t^L(m,0)-V_t^L((m-1)^+,0)$. Due to $\Delta V_t^L(m,0)\geq 0$ for all $L\geq m\geq 0$, this last inequality is satisfied. We now prove (6.7.6) for the case in which active action is optimal in both m and $m-1$, then, Inequality (6.7.6) reduces to

$$\begin{aligned}&\tilde{c}m+c_s+\lambda V_t^L(1,1)+\theta L V_t^L(0,1)+\mu V_t^L(0,0)\\&\geq\tilde{c}(m-1)+c_s+\lambda V_t^L(1,1)+\theta L V_t^L(0,1)+\mu V_t^L(0,0),\end{aligned}$$

which simplifies to $\tilde{c}\geq 0$, and hence (6.7.6) is satisfied. Let us now proceed with assuming that passive action is optimal in m and active in $m-1$, then the following holds,

$$\begin{aligned}&\tilde{c}m+\lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0)+\theta m V_t^L(m-1,0)+\left(\lambda\frac{m}{L}+\mu+\theta(L-m)\right)V_t^L(m,0)\\&\geq\tilde{c}(m-1)+\min\biggl(\lambda\left(1-\frac{m-1}{L}\right)V_t^L(m,0)+\left(\lambda\frac{m-1}{L}+\mu+\theta(L-m+1)\right)V_t^L(m-1,0)\\&\geq\tilde{c}(m-1)+c_s+\lambda V_t^L(1,1)+\theta L V_t^L(0,1)+\mu V_t^L(0,0),\end{aligned}$$

where the first inequality has been proven above, and the second inequality follows from the fact that active action is optimal in m. Therefore, Inequality (6.7.6) holds in case active is optimal in $m-1$ and passive in m. We are left with the proof in the case where the optimal action is active in m and passive in $m-1$. We have

$$\tilde{c}m + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0) \geq \tilde{c}(m-1) + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0)$$
$$\geq \tilde{c}(m-1) + \min\Bigg(\lambda\left(1 - \frac{m-1}{L}\right) V_t^L(m,0) + \left(\lambda\frac{m-1}{L} + \mu + \theta(L-m+1)\right) V_t^L(m-1,0)$$
$$+ \theta(m-1)V_t^L((m-2)^+,0),$$

where the first inequality has been proven above and the second inequality follows from the fact that passive action in optimal in state $m-1$. Therefore, Inequality (6.7.6) holds.

We have therefore proven that for any t, $V_t(m,0)$ is non-decreasing in m. And $V_t^L(m,0) \to V^L(m,0)$ point-wise, then, $V^L(m,0)$ is non-decreasing.

Let us now prove $f_t^L(m,0) - f_t^L(m,1) \leq f_t^L(m',0) - f_t^L(m',1)$ where

$$f_t^L(m,0) = \lambda\left(1 - \frac{m}{L}\right) V_t^L(m+1,0) + \theta m V_t^L(m-1,0) + \left(\lambda\frac{m}{L} + \mu + \theta(L-m)\right) V_t^L(m,0),$$
$$f_t^L(m,1) = \tilde{c}m + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0).$$

Proof for $f_t^L(m,0) - f_t^L(m,1) \leq f_t^L(m',0) - f_t^L(m',1)$ for all $m' \geq m$. Substituting the expression of $f_t^L(m,a)$ the inequality reduces to

$$\tilde{c}(m'-m) \leq \tilde{c}(m'-m) + \lambda\left(1 - \frac{m'}{L}\right) V_t^L(m'+1,0)$$
$$+ \left(\lambda\frac{m'}{L} + \mu + \theta(L-m')\right) V_t^L(m',0) + \theta m' V_t^L(m'-1,0)$$
$$- \left(\lambda\left(1 - \frac{m}{L}\right) V_t^L(m+1,0) + \left(\lambda\frac{m}{L} + \mu + \theta(L-m)\right) V_t^L(m,0) + \theta m V_t^L(m-1,0)\right).$$

Let us now define $m' = m + u$, with $u \geq 1$, and then the latter inequality writes

$$0 \leq \lambda\left(1 - \frac{m'}{L}\right)(V_t^L(m'+1,0) - V_t^L(m+1,0)) - \lambda\frac{u}{L}(V_t^L(m+1,0) - V_t^L(m,0))$$
$$+ \lambda\frac{m'}{L}(V_t^L(m',0) - V_t^L(m,0)) + \mu(V_t^L(m',0) - V_t^L(m,0)) + \theta m(V_t^L(m'-1,0) - V_t^L(m-1,0))$$
$$+ \theta u(V_t^L(m'-1,0) - V_t^L(m,0)),$$

which is satisfied due to $V_t^L(m,0)$ being non-decreasing.

6.7.2 Proof of Proposition 6.2

First we derive the expression of $\pi(m,1)$ and subsequently the expression of $\pi(m,0)$.

Steady-state distribution in the busy period

We first define the ordinary generating function that corresponds to $\pi(m,1)$ for all $m \in \mathbb{N}_0$, that is, $\Pi_1(z) = \sum_{m=0}^{\infty} z^m \pi(m,1)$, and recall Equation (6.4.2) for all $m \in \mathbb{N}$. Then, upon multiplying Equation (6.4.2) for state m with z^m, namely

$$z^m(\lambda + m\theta + \mu)\pi(m,1) = z^m\lambda\pi(m-1,1) + z^m(m+1)\theta\pi(m+1,1), \quad \forall\, m \in \mathbb{Z}\backslash\{0\},$$

and summing up the latter for all $m \in \{1, 2, \ldots\}$, we obtain

$$(\lambda+\mu)\left(\Pi_1(z) - \pi(0,1)\right) + \theta z \frac{\mathrm{d}}{\mathrm{d}z}\Pi_1(z) = \lambda z \Pi_1(z) + \theta\left(\frac{\mathrm{d}}{\mathrm{d}z}\Pi_1(z) - \pi(1,1)\right).$$

After some algebra the latter reduces to

$$\frac{(\lambda(1-z)+\mu)\Pi_1(z)}{-\theta(1-z)} + \frac{\mathrm{d}\Pi_1(z)}{\mathrm{d}z} = \frac{(\lambda+\mu)\pi(0,1) - \theta\pi(1,1)}{-\theta(1-z)}. \tag{6.7.7}$$

We now solve this ordinary differential equation. To do so, let us define $\Pi_1^n(z) = f_1(z)g_1(z)$ such that

$$\frac{\frac{\mathrm{d}f_1(z)}{\mathrm{d}z}}{f_1(z)} = -\frac{\lambda(1-z)+\mu}{-\theta(1-z)} \Rightarrow f_1(z) = \mathrm{e}^{\frac{\lambda z}{\theta}}(1-z)^{-\frac{\mu}{\theta}}. \tag{6.7.8}$$

Substituting

$$\Pi_1(z) = f_1(z)g_1(z) = \frac{\mathrm{e}^{\lambda z/\theta}}{(1-z)^{\mu/\theta}} g_1(z),$$

in Equation (6.7.7) and dividing both sides of the equality by $-\theta(1-z)\frac{\mathrm{e}^{\lambda z/\theta}}{(1-z)^{\mu/\theta}}$, we obtain

$$\frac{\mathrm{d}g_1(z)}{\mathrm{d}z} = \frac{(\lambda+\mu)\pi(0,1) - \theta\pi(1,1)}{-\theta(1-z)\mathrm{e}^{\lambda z/\theta}(1-z)^{-\mu/\theta}}.$$

By integrating this last equation, and noting that, since $f_1(0) = 1$ and $\Pi_1(0) = \pi(0,1)$, then $g_1(0) = \pi(0,1) \neq 0$, we derive

$$\begin{aligned} g_1(z) &= \pi(0,1) - \int_0^z \frac{(\lambda+\mu)\pi(0,1) - \theta\pi(1,1)}{\theta \mathrm{e}^{\lambda x/\theta}(1-x)^{1-\mu/\theta}} \mathrm{d}x \\ \Longrightarrow g_1(z) &= \pi(0,1) - \frac{(\lambda+\mu)\pi(0,1) - \theta\pi(1,1)}{\theta} \int_0^z \frac{(1-x)^{\mu/\theta}}{\mathrm{e}^{\lambda x/\theta}(1-x)} \mathrm{d}x. \end{aligned} \tag{6.7.9}$$

We now aim at deriving an explicit expression for $\pi(m,1) = \frac{1}{m!}\frac{\mathrm{d}^m\Pi_1(z)}{\mathrm{d}z^m}|_{z=0}$ for all $m \geq 0$. From (6.7.8) and (6.7.9) we have

$$\Pi_1(z) = \frac{\pi(0,1)\mathrm{e}^{\frac{\lambda z}{\theta}}}{(1-z)^{\frac{\mu}{\theta}}} - \frac{((\lambda+\mu)\pi(0,1) - \theta\pi(1,1))}{\theta(1-z)^{\frac{\mu}{\theta}}} \mathrm{e}^{\frac{\lambda}{\theta}(z-1)} \cdot (-1)^{\frac{\mu}{\theta}-1}\left(\frac{\theta}{\lambda}\right)^{\frac{\mu}{\theta}} \int_{-\frac{\lambda}{\theta}}^{-\frac{\lambda}{\theta}(1-z)} y^{\frac{\mu}{\theta}-1}\mathrm{e}^{-y}\mathrm{d}y, \tag{6.7.10}$$

where we have used a change of variable $y = -\frac{\lambda}{\theta}(1-x)$ in the integral. Observe that the integral that shows up in the expression of $\Pi_1(z)$ is an incomplete gamma function Abramowitz *et al.* [2, Chap. 6].

Therefore, since $\mu/\theta > 0$,

$$\int_{-\frac{\lambda}{\theta}}^{-\frac{\lambda}{\theta}(1-z)} y^{\frac{\mu}{\theta}-1}\mathrm{e}^{-y}\mathrm{d}y = \left(-\frac{\lambda}{\theta}\right)^{\frac{\mu}{\theta}} \sum_{i=0}^{\infty} \frac{(\frac{\lambda}{\theta})^i((1-z)^{i+\frac{\mu}{\theta}}-1)}{i!(\frac{\mu}{\theta}+i)}. \tag{6.7.11}$$

Before deriving the probabilities $\pi(m,1)$ for all $m \in \mathbb{N}_0$, note that $\Pi_1(z)$ is not well defined in $z=1$, and therefore we force $\lim_{z\to 1}\Pi_1(z)$ to be a $0/0$ type of indeterminate. By letting the limit as $z \to 1$ of the numerator in (6.7.10) be 0 we obtain the condition

$$\pi(0,1)\mathrm{e}^{\lambda/\theta} - \left(\frac{\lambda+\mu}{\theta}\pi(0,1) - \pi(1,1)\right)\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)} = 0.$$

Solving the latter equation we obtain the explicit expression of $\pi(1,1)$ with respect to $\pi(0,1)$, namely,

$$\pi(1,1) = a_1\pi(0,1), \quad a_1 = \frac{\lambda+\mu}{\theta} - \frac{\mathrm{e}^{\lambda/\theta}}{\sum_{i=0}^{\infty}\frac{(\lambda/\theta)^i}{i!(\mu/\theta+i)}}.$$

After substituting $\pi(1,1) = a_1\pi(0,1)$ and (6.7.11) in (6.7.9), from (6.7.8) and (6.7.9) we obtain

$$\begin{aligned} f_1(z) &= \mathrm{e}^{\frac{\lambda z}{\theta}}(1-z)^{-\frac{\mu}{\theta}}, \\ g_1(z) &= \pi(0,1)\left(1 + \frac{\sum_{j=0}^{\infty}\frac{(\frac{\lambda}{\theta})^j((1-z)^{j+\frac{\mu}{\theta}}-1)}{j!(\frac{\mu}{\theta}+j)}}{\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)}}\right). \end{aligned}$$

We can now proceed to compute the steady-state distribution in the busy period, that is, $\pi(m,1)$ for all $m \geq 1$. Let us define $\ell_i(\mu/\theta) = \mu/\theta \cdot \ldots \cdot (\mu/\theta + i - 1)$ for all $i \geq 1$ and $\ell_0(\mu/\theta) = 1$ and note that $\pi(m,1) = \frac{1}{m!}\frac{\mathrm{d}^m\Pi_1(z)}{\mathrm{d}z^m}|_{z=0} = \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}f_1^{(m-k)}g_1^{(k)}$, where

$$f_1^{(k)} := \frac{\mathrm{d}^k f_1(z)}{\mathrm{d}z^k}|_{z=0} = \sum_{i=0}^{k}\binom{k}{i}\left(\frac{\lambda}{\theta}\right)^{k-i}\frac{\mathrm{e}^{\frac{\lambda z}{\theta}}\ell_i\left(\frac{\mu}{\theta}\right)}{(1-z)^{\frac{\mu}{\theta}+i}}|_{z=0} = \sum_{i=0}^{k}\binom{k}{i}\left(\frac{\lambda}{\theta}\right)^{k-i}\ell_i\left(\frac{\mu}{\theta}\right), \text{ for all } k \geq 0,$$

$$g_1^{(k)} := \frac{\mathrm{d}^k g_1(z)}{\mathrm{d}z^k}|_{z=0} = \pi(0,1)\frac{\sum_{j=0}^{\infty}\frac{(\frac{\lambda}{\theta})^j\ell_k(-\frac{\mu}{\theta}-j)(1-z)^{j+\frac{\mu}{\theta}-k}}{j!(\frac{\mu}{\theta}+j)}}{\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)}}|_{z=0} = \pi(0,1)\frac{\sum_{j=0}^{\infty}\frac{(\frac{\lambda}{\theta})^j\ell_k(-\frac{\mu}{\theta}-j)}{j!(\frac{\mu}{\theta}+j)}}{\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)}}, \text{for all } k \geq 1,$$

and $g_1^{(0)} = \pi(0,1)$. Define $a_1(0) := 1, a_1(1) := a_1$ and $a_1(m) := \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}f_1^{(m-k)}g_1^{(k)}$ for all $m \geq 2$. Then we obtain $\pi(m,1) = a_1(m)\pi(0,1)$, with $a_1(m)$ given as in Proposition 6.2.

Steady-state distribution in the idle period

We first define the ordinary generating function that corresponds to $\pi(m,0)$ for all $0 \leq m \leq n-1$, that is, $\Pi_0(z) = \sum_{m=0}^{\infty}z^m\pi(m,0) = \sum_{m=0}^{n-1}z^m\pi(m,0)$, where by definition $\pi(m,0) = 0$ for all $m \geq n$, and recall Equation (6.4.3) for all $1 \leq m \leq n-1$. Upon multiplying Equation (6.4.3) in state m with z^m, namely

$$z^m(\lambda+m\theta)\pi(m,0) = z^m\lambda\pi(m-1,0) + z^m\mu\pi(m,1) + z^m(m+1)\theta\pi(m+1,0),$$

and summing the latter over all $1 \leq m \leq n-1$, we then obtain

$$\lambda(\Pi_0(z) - \pi(0,0)) + \theta z \frac{\mathrm{d}\Pi_0(z)}{\mathrm{d}z} = \theta\left(\frac{\mathrm{d}\Pi_0(z)}{\mathrm{d}z} - \pi(1,0)\right) + \lambda z(\Pi_0(z) - \pi(n-1,0)) + \mu \sum_{m=1}^{n-1} z^m \pi(m,1).$$

Using (6.4.3) in the case $m = 0$, that is, $\lambda\pi(0,0) - \theta\pi(1,0) = \mu\pi(0,1)$, and after some algebra, we derive

$$-\frac{\lambda}{\theta}\Pi_0(z) + \frac{\mathrm{d}\Pi_0(z)}{\mathrm{d}z} = \frac{\lambda z}{\theta(1-z)}\pi(n-1,0) - \frac{\mu}{\theta(1-z)} \sum_{m=0}^{n-1} z^m \pi(m,1). \tag{6.7.12}$$

Observe in the latter equation that for $\Pi_0(z)$ to be well defined in $z = 1$, which we know equals $\sum_{m=0}^{n-1} \pi(m,0) < 1$, the condition

$$\lim_{z \to 1} \lambda z \pi(n-1,0) - \mu \sum_{m=0}^{n-1} \pi(m,1) = 0,$$

needs to be satisfied. We then obtain the extra condition to the problem

$$\pi(n-1,0) = \frac{\mu}{\lambda} \sum_{m=0}^{n-1} \pi(m,1) = \frac{\mu}{\lambda}\pi(0,1) \sum_{m=0}^{n-1} a_1(m),$$

with $a_1(m)$ as given by Proposition 6.2. This yields $\pi(n-1,0) = a_0^n(n-1)\pi(0,1)$. To derive the expression of $\pi(m,0)$ for all $n-2 \geq m \geq 1$, we adopt the following balance equations for all $n-2 \geq m \geq 1$, which are equivalent to those introduced in Equation (6.4.3):

$$\lambda\pi(m,0) = \theta(m+1)\pi(m+1,0) + \lambda\pi(n-1,0) - \mu \sum_{j=m+1}^{n-1} \pi(j,1).$$

This equation can be solved using similar arguments as those used in Section 6.4.1, since the first two terms on the RHS of the equation correspond to the balance equations for the case $\mu = \infty$. Having noticed this, the recursion can easily be solved to obtain $a_0^n(m)$ for all $n-2 \geq m \geq 1$ as given in Proposition 6.2. To do so recall Equation (6.4.3) and note that one can equivalently write for all $n-2 \geq m \geq 1$

$$\lambda\pi(m,0) = \theta(m+1)\pi(m+1,0) + \lambda\pi(n-1,0) - \mu \sum_{j=m+1}^{n-1} \pi(j,1).$$

Then,

$$\begin{aligned}
\pi(m,0) =& \frac{\theta(m+1)}{\lambda}\pi(m+1,0)+\pi(n-1,0)-\frac{\mu}{\lambda}\sum_{j=m+1}^{n-1}\pi(j,1),\\
=& \frac{\theta(m+1)}{\lambda}\left(\frac{\theta(m+2)}{\lambda}\pi(m+2,0)+\pi(n-1,0)-\frac{\mu}{\lambda}\sum_{j=m+2}^{n-1}\pi(j,1)\right)+\pi(n-1,0)-\frac{\mu}{\lambda}\sum_{j=m+1}^{n-1}\pi(j,1)\\
=& \frac{\theta^2}{\lambda^2}(m+1)(m+2)\pi(m+2,0)+\left(1+\frac{\theta}{\lambda}(m+1)\right)\pi(n-1,0)-\frac{\mu}{\lambda}\pi(m+1,1)\\
& -\frac{\mu}{\lambda}\sum_{j=m+2}^{n-1}\left(1+\frac{\theta(m+1)}{\lambda}\right)\pi(j,1)\\
=& \dots\\
=& \pi(n-1,0)\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}-\frac{\mu}{\lambda}\sum_{j=m+1}^{n-1}\pi(r,1)\sum_{i=0}^{r-m-1}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}.
\end{aligned}$$

This last expression is valid for all $m \geq 1$.

Finally, the expression for $a_0^n(0)$ can be derived by solving $\pi(0,0)=\frac{\mu}{\lambda}\pi(0,1)+\frac{\theta}{\lambda}a_0^n(1)\pi(0,1)$. Then, we obtain $\pi(m,0)=a_0^n(m)\pi(0,1)$.

6.7.3 Proof of Proposition 6.3

We aim at proving that for all $n' \neq n$

$$\tilde{c}\mathbb{E}(N^n)+c_sP_b^n \leq \tilde{c}\mathbb{E}(N^{n'})+c_sP_b^{n'}.$$

We present here the proof in the case $n' < n$, the other case can be done similarly. By assumption we have for all $n \geq 1$, $\alpha(n-1) \leq \alpha(n)$, then

$$\begin{aligned}
&\frac{\mathbb{E}(N^{n-1})-\mathbb{E}(N^{n-2})}{P_b^{n-2}-P_b^{n-1}} \leq \frac{\mathbb{E}(N^n)-\mathbb{E}(N^{n-1})}{P_b^{n-1}-P_b^n}\\
&\Longrightarrow (\mathbb{E}(N^{n-1})-\mathbb{E}(N^{n-2}))(P_b^{n-1}-P_b^n) \leq (\mathbb{E}(N^n)-\mathbb{E}(N^{n-1}))(P_b^{n-2}-P_b^{n-1}).
\end{aligned}$$

In the latter inequality we sum and subtract $\mathbb{E}(N^n)(P_b^{n-1}-P_b^n)$ on the left hand side, that is,

$$(\mathbb{E}(N^{n-1})-\mathbb{E}(N^n)+\mathbb{E}(N^n)-\mathbb{E}(N^{n-2}))(P_b^{n-1}-P_b^n) \leq (\mathbb{E}(N^n)-\mathbb{E}(N^{n-1}))(P_b^{n-2}-P_b^{n-1}),$$

after some algebra this last inequality reduces to

$$\frac{\mathbb{E}(N^n)-\mathbb{E}(N^{n-2})}{P_b^{n-2}-P_b^n} \leq \frac{\mathbb{E}(N^n)-\mathbb{E}(N^{n-1})}{P_b^{n-1}-P_b^n} \leq c_s.$$

Similarly, one can prove

$$\alpha(n-1) \leq \frac{\mathbb{E}(N^n)-\mathbb{E}(N^{n-2})}{P_b^{n-2}-P_b^n}.$$

We now make the following induction assumption for a given $a > 2$

$$\alpha(n-a+1) \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a})}{P_b^{n-a} - P_b^n} \leq c_s.$$

It has been proven above for $a = 1, 2$. By assumption on the statement we have $\alpha(n-a) \leq \alpha(n-a+1)$, hence from the latter equation we obtain

$$\begin{aligned}&\alpha(n-a) \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a})}{P_b^{n-a} - P_b^n}\\ &\implies (\mathbb{E}(N^{n-a}) - \mathbb{E}(N^{n-a-1}))(P_b^{n-a} - P_b^n) \leq (\mathbb{E}(N^n) - \mathbb{E}(N^{n-a}))(P_b^{n-a-1} - P_b^{n-a}),\end{aligned}$$

adding and subtracting $\mathbb{E}(N^n)(P_b^{n-a} - P_b^n)$ on the left hand side, and after some algebra, we obtain

$$\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a-1})}{P_b^{n-a-1} - P_b^n} \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a})}{P_b^{n-a} - P_b^n} \leq c_s. \tag{6.7.13}$$

From the latter we observe that

$$(\mathbb{E}(N^n) - \mathbb{E}(N^{n-a-1}))(P_b^{n-a} - P_b^n + P_b^{n-a-1} - P_b^{n-a-1}) \leq (\mathbb{E}(N^n) - \mathbb{E}(N^{n-a}))(P_b^{n-a-1} - P_b^n),$$

which after some algebra reduces to

$$\alpha(n-a) \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a-1})}{P_b^{n-a-1} - P_b^n}.$$

The latter together with (6.7.13) gives

$$\alpha(n-a) \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a-1})}{P_b^{n-a-1} - P_b^n} \leq c_s,$$

which concludes the induction. For all $0 \leq a \leq n-1$ denote $n' = n-1-a$. We have proven that for all $n' < n$

$$\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n'})}{P_b^{n'} - P_b^n} \leq c_s \implies \tilde{c}\mathbb{E}(N^n) + c_s P_b^n \leq \tilde{c}\mathbb{E}(N^{n'}) + c_s P_b^{n'}.$$

Which concludes the proof.

6.7.4 Proof of Lemma 6.1

Let us first proof that P_b^n is non-increasing. It suffices to prove $\pi_{n-1}^n \leq \pi_{n-2}^{n-1}$ for all $n \geq 2$. This inequality writes

$$\sum_{m=0}^{n-2} \sum_{i=0}^{n-2-m} \left(\frac{\theta}{\lambda}\right)^i \frac{(m+i)!}{m!} \leq \sum_{m=0}^{n-1} \sum_{i=0}^{n-1-m} \left(\frac{\theta}{\lambda}\right)^i \frac{(m+i)!}{m!} \Leftrightarrow 0 \leq \sum_{m=0}^{n-1} \left(\frac{\theta}{\lambda}\right)^{n-1-m} \frac{(n-1)!}{m!}.$$

The RHS is positive for all $n \geq 2$ and hence $\pi_{n-1}^{n} \leq \pi_{n-2}^{n-1}$. Having proven that π_{n-1}^{n} is non-increasing in n, we now proceed to prove the convexity of P_b^n. Convexity of P_b^n is implied by $\pi_{n-1}^{n} - \pi_{n}^{n+1} \leq \pi_{n-2}^{n-1} - \pi_{n-1}^{n}$, which after substitution of the corresponding values reduces to

$$\frac{\sum_{m=0}^{n+1}\sum_{i=0}^{n+1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!} - \sum_{m=0}^{n}\sum_{i=0}^{n-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}}{\sum_{m=0}^{n+1}\sum_{i=0}^{n+1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}}$$
$$\leq \frac{\sum_{m=0}^{n}\sum_{i=0}^{n-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!} - \sum_{m=0}^{n-1}\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}}{\sum_{m=0}^{n-1}\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}}. \tag{6.7.14}$$

After some algebra the latter reduces to

$$\left(\sum_{m=0}^{n+1}\left(\frac{\theta}{\lambda}\right)^{n+1-m}\frac{(n+1)!}{m!}\right)\sum_{m=0}^{n-1}\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}$$
$$\leq \left(\sum_{m=0}^{n}\left(\frac{\theta}{\lambda}\right)^{n-m}\frac{n!}{m!}\right)\sum_{m=0}^{n+1}\sum_{i=0}^{n+1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!},$$

We have not been able to prove the latter formally, however, we have computed the latter inequality numerically for increasing values of n and random values of λ and θ for which we have seen that it holds. Later we have analytically computed the latter as n goes to infinity and the inequality holds as well. Hence, P_b^n is convex in n.

Let us now prove that $\alpha(n)$ as defined in Proposition 6.3 is non-decreasing, that is,

$$\alpha(n) := \tilde{c}\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-1})}{P_b^{n-1} - P_b^n} = \frac{\sum_{m=0}^{n-1} m\pi_m^n - \sum_{m=0}^{n-2} m\pi_m^{n-1}}{P_b^{n-1} - P_b^n}$$
$$= \frac{\sum_{m=0}^{n-1} m\left(\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\right)\pi_{n-1}^n}{P_b^{n-1} - P_b^n} - \frac{\sum_{m=0}^{n-2} m\left(\sum_{i=0}^{n-2-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\right)\pi_{n-2}^{n-1}}{P_b^{n-1} - P_b^n}.$$

The latter after some algebra writes

$$\alpha(n) = \frac{\sum_{m=0}^{n-1} m\left(\frac{\theta}{\lambda}\right)^{n-1-m}\frac{(n-1)!}{m!}\pi_{n-1}^n}{P_b^{n-1} - P_b^n} + \sum_{m=0}^{n-2} m\sum_{i=0}^{n-2-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\left(\frac{\pi_{n-1}^n - \pi_{n-2}^{n-1}}{P_b^{n-1} - P_b^n}\right).$$

We now aim at proving $\alpha(n) \leq \alpha(n+1)$ for all n, that is if

$$\frac{\sum_{m=0}^{n-1} m\left(\frac{\theta}{\lambda}\right)^{n-1-m}\frac{(n-1)!}{m!}\pi_{n-1}^n}{P_b^{n-1} - P_b^n} + \sum_{m=0}^{n-2} m\sum_{i=0}^{n-2-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\left(\frac{\pi_{n-1}^n - \pi_{n-2}^{n-1}}{P_b^{n-1} - P_b^n}\right)$$
$$\leq \frac{\sum_{m=0}^{n} m\left(\frac{\theta}{\lambda}\right)^{n-m}\frac{n!}{m!}\pi_{n}^{n+1}}{P_b^{n} - P_b^{n+1}} + \sum_{m=0}^{n-1} m\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\left(\frac{\pi_{n}^{n+1} - \pi_{n-1}^{n}}{P_b^{n} - P_b^{n+1}}\right),$$

The second term in the LHS of the inequality being less than or equal to the second term in the RHS of the inequality follows from P_b^n and π_{n-1}^n being convex non-increasing functions in n. We are therefore left to proof

$$\sum_{m=0}^{n-1} m \left(\frac{\theta}{\lambda}\right)^{n-1-m} \frac{(n-1)!}{m!}\pi_{n-1}^{n} \le \sum_{m=0}^{n} m \left(\frac{\theta}{\lambda}\right)^{n-m} \frac{n!}{m!}\pi_{n}^{n+1}.$$

We have not been able to prove the latter formally, however, we have computed the latter inequality numerically for increasing values of n and random values of λ and θ for which we have seen that it holds. Later we have analytically computed the latter as n goes to infinity and the inequality holds as well.

6.7.5 Justification of Conjecture 6.1

To prove that P_b^n is non-increasing it suffices to prove $\pi^n(0,1) \ge \pi^{n+1}(0,1)$. Observe that $\pi^n(0,1) \ge \pi^{n+1}(0,1)$ is implied by $\sum_{m=0}^{n-1} a_0^n(m) \le \sum_{m=0}^{n} a_0^{n+1}(m)$. Furthermore, after some algebra one can obtain that

$$\sum_{m=0}^{n} a_0^{n+1}(m) = \sum_{m=0}^{n} a_0^{n}(m) + \frac{\mu}{\lambda}\sum_{r=0}^{n} a_1(r)\left(\frac{\theta}{\lambda}\right)^{n-m} \frac{n!}{m!},$$

therefore

$$0 \le \sum_{m=0}^{n} a_0^{n+1}(m) - \sum_{m=0}^{n-1} a_0^{n}(m) \iff 0 \le a_0^n(n) + \frac{\mu}{\lambda}\sum_{r=0}^{n} a_1(r)\left(\frac{\theta}{\lambda}\right)^{n-m} \frac{n!}{m!}.$$

The last inequality is satisfied since $a_0^n(m) > 0$ for all m.

Appendix A

Appendix

In this appendix we present a compilation of the most relevant results regarding optimality of stochastic and deterministic control problems that have been used throughout the thesis. In Section A.1 we present the conditions for the existence of an optimal stationary solution for the average cost criteria. In Section A.2 we give Pontryagin's Minimum Principle, that is, necessary conditions for optimality in a deterministic optimal control problem. In Section A.3 we present sufficient conditions for optimality in a deterministic optimal control problem.

A.1 Conditions for the existence of optimal stationary policies: The Bellman equation

We consider a continuous time MDP with the cost per unit of time given by $C(m, a)$ and the transition rates $q^a(m, \tilde{m})$ for all $m, \tilde{m} \in E$ and $a \in \mathcal{A}$, where E is the state space and $\mathcal{A}$ the action space. We assume $E \subseteq \mathbb{N}^d$, for some $d \in \mathbb{N}$ and $\mathcal{A}$ compact. Recall the objective function in (1.3.4), and define

$$\mathcal{C}^\phi = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T C(N^\phi(t), S^\phi(N^\phi(t)))\mathrm{d}t\right).$$

The average reward optimality equation for this continuous-time MDP can then be written as

$$\tilde{g} + \tau(m,a)V(m) = \min_{a\in\mathcal{A}}\{C(m,a) + \sum_{\tilde{m}\in E} q^a(m,\tilde{m})V(\tilde{m})\}. \tag{A.1.1}$$

where $\tilde{g}$ is the optimal average reward per unit of time. The latter when the system is uniformizable (*i.e.*, $\tau(m,a) \leq b$ for all states $m \in E$ and all actions $a \in \mathcal{A}$) can be transformed into a discrete-time MDP and might be solved using the techniques that are well known for discrete-time MDPs. The optimality results for a discrete-time MDP read as follows.

Theorem A.1. *If there exist a function $V(\cdot)$ and $g \in \mathbb{R}$ such that*

$$g + V(m) = \min_{a\in\mathcal{A}}\{\tilde{C}(m,a) + \sum_{\tilde{m}\in E} p^a(\tilde{m},m)V(\tilde{m})\},$$

where $\tilde{C}(m,a) = C(m,a)/b$, $p^a(m,\tilde{m}) = q^a(m,\tilde{m})/b$ *and* $g = \tilde{g}/b$. *Then there exists a stationary policy* ϕ^* *such that*

$$g = \mathcal{C}^{\phi^*} = \min_{\phi} \mathcal{C}^{\phi}.$$

For additional information and the proof of Theorem A.1 (and the required conditions) we refer to Guo *et al.* [54, Chapter 2.2], Puterman [77, Chapter 11.4] and Ross [80, Chapter V.2].

A.2 Necessary conditions: Pontryagin's Minimum Principle

Consider the following time-invariant optimal control problem. The objective is to find a control trajectory $s(t)$ for all $t \in [0,T]$ together with the state trajectory $m(t)$ such that the functional

$$\int_0^T C(m(t), s(t))\,\mathrm{d}t, \tag{A.2.1}$$

is minimized, where the cost function $C(\cdot,\cdot)$ is continuously differentiable with respect to $m(\cdot)$, continuous with respect to $s(\cdot)$ and T is the terminal time which is subject to optimization. We further assume that $m(t) \in \mathbb{R}^d$, for some $d \in \mathbb{N}$, and $s(t) \in \mathcal{S}$.

The dynamics are given by a set of differential equations, that is,

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} = f(m(t), s(t)), \text{ for all } t \in [0,T],$$

and the initial state is $m(0) = m_0$. Additionally we consider that

$$h_1(s(t)) \leq 0, \text{ and } h_2(m(t)) \leq 0, \text{ for all } t \in [0,T], \tag{A.2.2}$$

where $h_1(\cdot)$ is a pure control constraint, and $h_2(\cdot)$ is a pure state constraint. We are also going to assume that the following qualification conditions hold. The matrices

$$\begin{bmatrix} \frac{\partial h_1(s(t))}{\partial s(t)} & \operatorname{diag}(h_1) \end{bmatrix}, \text{ and } \begin{bmatrix} \frac{\partial h_2^1(m(t))}{\partial s(t)} \end{bmatrix},$$

have full rank, where $h_2^1 = \frac{\partial h_2(m(t))}{\partial m(t)} f(m(t), s(t))$.

Recall the definition of the Hamiltonian and the Lagrangian introduced in Section 1.3.4, namely,

$$\mathcal{H}(m(t), s(t), \gamma(t)) := C(m(t), s(t)) + \gamma^{\mathsf{T}}(t) f(m(t), s(t)), \text{ and,}$$

$$\mathcal{L}(m(t), s(t), \gamma(t), \nu(t), \omega(t)) := \mathcal{H}(m(t), s(t), \gamma(t)) + \nu^{\mathsf{T}}(t) h_1(s(t)) + \omega^{\mathsf{T}}(t) h_2(m(t)),$$

where $\omega(\cdot)$ and $\nu(\cdot)$ are the Lagrange multipliers and $\gamma(\cdot)$ the adjoint vector. And we denote by $\mathcal{H}(t) := \mathcal{H}(m(t), s(t), \gamma(t))$ and $\mathcal{L}(t) := \mathcal{L}(m(t), s(t), \gamma(t), \nu(t), \omega(t))$. As mentioned in the introductory chapter of the thesis, necessary conditions for optimality can be used to find candidates for optimality. Sufficient conditions can then be used to find the optimal solution among these candidates, see [84, Chapter 6].

Another possibility is to compare all candidates to find the optimal one. The following theorem, given that the qualification conditions hold, provides necessary conditions for the optimality of a $(m(t), s(t))$ pair. We refer to the survey paper Hartl *et al.* [56] for an overview on more general minimum principles.

Theorem A.2. *Let $s^*(\cdot)$ be the optimal control, piecewise continuous, let T be the optimal final time (subject to optimization) and $m^*(\cdot)$ the associated optimal trajectory defined in the interval $[0, T]$. Then, there exists a piecewise continuous adjoint function $\gamma^*(t) = (\gamma_1^*(t), \ldots, \gamma_K^*(t))$ with piecewise continuous derivatives, piecewise continuous multipliers $\nu(t)$ and $\omega(t)$ for all $t \in [0, T]$ that satisfy,*

1. $$\dot{\gamma}^* = -\frac{\partial \mathcal{L}(t)}{\partial m^*}, \tag{A.2.3}$$

   in all the continuity points of $m^*(t)$ where $\mathcal{L}$ is the Lagrangian of the system.

2. $$s^*(t) = \arg\min_{s(t)\in\mathcal{S}} \mathcal{H}(m^*(t), s^*(t), \gamma^*(t)), \tag{A.2.4}$$

   with $\mathcal{U}$, the set of admissible controls. Then, the latter condition writes

   $$\frac{\partial \mathcal{L}^*(t)}{\partial s^*} = 0, \tag{A.2.5}$$

3. $$\dot{m}^*(t) = f(m^*(t), s^*(t)), \tag{A.2.6}$$

   in all the continuity points of $s^*(t)$, with $m^*(0) = m_0$.

4. *Also*

   $$\nu(t)h_1(s^*(t)) = 0, \nu(t) \geq 0 \text{ and } \omega(t)h_2(m^*(t)) = 0, \omega(t) \geq 0 \text{ for all } t \in [0, T]. \tag{A.2.7}$$

   Furthermore, there exists a constant K_0 that satisfies

   $$\mathcal{H}(m^*(t), s^*(t), \gamma^*(t)) = K_0, \text{ for all } t \in [0, T]. \tag{A.2.8}$$

The constant K_0 for time-invariant free final time problems equals 0.

A.3 Sufficient conditions: Hamilton-Jacobi-Bellman Equation

In this section we present sufficient conditions for optimality for a time invariant infinite horizon optimal control problem. That is, we consider the problem where the objective is to find a control trajectory $s(t)$ for all $t \in [0, \infty)$ together with the state trajectory $m(t)$ such that the functional

$$\int_0^\infty C(m(t), s(t))\, \mathrm{d}t, \tag{A.3.1}$$

is minimized, where the cost function $C(\cdot, \cdot)$ is continuously differentiable with respect to $m(\cdot)$ and continuous with respect to $s(\cdot)$. We further assume that $m(t) \in \mathbb{R}^d$, for some $d \in \mathbb{N}$, and $s(t) \in \mathcal{S}$.

The dynamics are given by a set of differential equations, that is,

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} = f(m(t), s(t)), \text{ for all } t,$$

and the initial state is $m(0) = m_0$. This conditions are given by the HJB equation, see Bertsekas [22, Prop. 3.2.1].

Theorem A.3. *Suppose there exists a function $V(t, m)$ continuously differentiable in t and m such that*

$$0 = \min_{s \in \mathcal{S}} [C(m, s) + \nabla_m V(t, m)^\intercal f(m, s)], \text{ for all } t, m. \tag{A.3.2}$$

If the control trajectory $s^*(t)$ minimizes the right hand side of equation (1.3.12) for every $t \in [0, T]$, $s^*(t)$ is piecewise continuous in t and $\frac{\mathrm{d}m(t)}{\mathrm{d}t} = f(m^*(t), s^*(t))$ has unique solution starting at any pair (t, m), with $m^*(t)$ the trajectory corresponding to $s^*(t)$. Then, $V(t, m)$ is the optimal cost for all t, m, and the control trajectory $s^*(x)$ for all $t \in [0, T]$ is optimal.

Résumé en français

Dans cette thèse nous ètudions le contrôle dynamique des systèmes de partage de ressources qui interviennent dans plusieurs domaines: *e.g.*, gestion d'inventaire, santé, réseaux de communications. De par ses applications variées, le problème de l'allocation de ressources pour des classes de clients hétérogènes et concurrentes attire fortement l'attention de la communauté dans le domaine. Une politique de partage de ressource peut-être conçue pour satisfaire un certain critère de performance: *e.g*, maximisation des ventes en gestion des stocks, minimisation du nombre de victimes dans un système de trie ou maximisation du débit dans un réseau. De nombreuses facettes de ces systèmes peuvent être capturées mathématiquement par des modèles théoriques de file d'attente. Dans cette thèse, c'est cette approche de modélisation que nous considérons. Les modèles théoriques ont été traditionnellement utilisés pour évaluer la performance des systèmes de file d'attente, *e.g.*, prédire la distribution de leur tailles, calculer des temps de séjours, mais aussi afin d'aider à la prise de décision, *e.g.*, développer une politique d'ordonnancement pour atteindre un objectif préalablement fixé. Ce dernier cas constituera le principal objet de cette thèse. Nombre de systèmes de partage de ressources possèdent une nature aléatoire et pour cette raison, nous avons considéré des modèles de contrôle stochastique. Plus spécifiquement, nous étudions un large spectre de systèmes de partage de ressources qui entrent dans le cadre des problèmes des bandits restless (RBPs de l'anglais restless bandit problems), *e.g*, service unique de files d'attentes multi-classes avec abandons, ordonnancement opportuniste dans les réseaux sans fil, et gestion d'inventaire en présence d'éléments périssables, etc. Un RBP, qui est une classe générale de problème d'optimisation stochastique dynamique, est décrit comme suit. Il existe K classes de bandits dans un système dont au plus M peuvent être activés, l'objectif est de choisir l'ensemble des bandits à activer dans le but de minimiser leur coût. Dans cette configuration, deux types de RBPs sont considérés. Ceux du choix optimal de la classe et ceux du partage optimal de charge. Dans les deux cas, les bandits représentent les classes de clients. Dans les problèmes de choix optimal de classe, l'activation d'un bandit implique qu'une ressource soit allouée à la classe qui lui est associée, tandis que dans les problèmes de partage optimal de charge un répartiteur décide à quelle classe doit appartenir un nouveau client. Dans ce dernier cas, l'activation d'un bandit implique qu'un nouveau client rejoindra la file qui correspond à sa classe. Dans cette thèse une attention particulière est donnée aux problèmes dans lesquels les clients peuvent abandonner le système avant d'avoir fini d'être servi. Les abandons ou renoncements sont présent lorsque les clients, insatisfait par leur long temps d'attente, décident volontairement de quitter le système. Ce phénomène a un impact considérable dans les applications d'Internet ou dans les centres d'appels, lorsque les clients peuvent abandonner en phase d'attente

comme en phase de service. l'abandon est un phénomène fortement indésirable, aussi bien du point de vue d'un client que de celui du système (un profit est perdu). Il peut avoir un impact économique considérable. Il n'est donc pas surprenant que ce sujet ait grandement attiré l'attention de la communauté scientifique, notamment ces dernières années. Pour certains problèmes considérés dans cette thèse, il est possible de déterminer un contrôle optimal. Cependant, dans la plupart des cas, trouver une solution optimale est inaccessible et il est nécessaire de trouver des heuristiques performantes. Ainsi, nous donnons une solution optimale lorsque ceci est possible et nous développons des heuristiques lorsqu'une solution optimale ne peut être trouvée. Une heuristique peut être obtenue en employant différentes techniques. Dans ces travaux, nous utiliserons principalement deux approches. La première, basée sur des techniques de contrôle stochastique sera exploitée en Partie I et en Partie III. La seconde repose sur des techniques de contrôle de fluide et fera l'objet de la partie II. Ces techniques sont introduites en détail dans le Chapitre 1, tout comme les exemples d'application considérés dans cette thèse. Dans la Partie I, nous nous focalisons sur les RBPs pour lesquels l'état des bandits (ou nombre de client en attente dans chaque classe) évolue suivant un processus de naissance et de mort. Nous considérons que K classes hétérogènes de clients sont en concurrence pour M ressources disponibles. Relaxer la contrainte sur le nombre de ressources, permet d'approcher le problème contraint d'optimisation original à K dimensions plus simplement, par K problèmes d'optimisation uni-dimensionel et sans contrainte. Cette méthode est aussi connue sous le nom de relaxation Lagrangienne. La solution au problème relaxé peut-être prouvée comme étant une politique basée sur des indices et peut être utilisable en tant qu'heuristique (politique de l'indice de Whittle) pour le problème originellement contraint. Une politique d'indice est telle qu'il existe une valeur (ou indice) attachée à chaque classe de clients (dépendant uniquement des paramètres de cette classe et de son état), pour lesquelles il est prescrit de servir la classe de client ayant l'indice le plus grand. Dans le Chapitre 2, nous dérivons une expression analytique pour l'indice de Whittle, comme une fonction de la distribution de probabilité des états stationnaires pour le cas d'un processus de naissance et de mort. Cette expression nécessite la vérification de plusieurs conditions techniques comme la propriété d'indexabilité, et de plus n'est calculable explicitement que dans certains cas spécifiques. Dans le Chapitre 3, nous appliquons les résultats obtenus dans le Chapitre 2 au cas particulier des abandons dans les files d'attentes multi-classes. Nous prouvons que les politiques monotones sont optimales pour le problème non-contraint et que les bandits sont indexables. Ces résultats fournissent une expression analytique de l'indice de Whittle qui permet dans certains cas une meilleure compréhension du système. Nous remarquons que les abandons peuvent rendre indexable des problèmes qui ne le sont initialement pas, en faisant tendre leur taux vers zéro. C'est par exemple le cas d'une file d'attente de type M/M/1 avec un coût d'attente convexe. Nous prouvons que l'indice de Whittle est asymptotique-optimal aussi bien en régime de fort traffic qu'en régime de faible traffic et nous évaluons numériquement la performance de l'indice de Whittle pour différent niveau de charge. Nous concluons que la politique de l'indice de Whittle est quasi-optimale en présence d'abandons. Dans la Partie II, nous dérivons des heuristiques en approchant des systèmes stochastiques de partage de ressources par des modèles déterministes de fluides. Deux approches différentes sont considérées. Dans le Chapitre 4, nous commençons par l'approche de relaxation telle que suivie dans la Partie I. Nous lissons la version relaxé du problème multi-classe et multi-serveur sous l'hypothèse que chaque file évolue selon un processus de naissance et de mort. Ceci soulève un problème non-contraint déterministe qui fourni une politique d'indice pour le problème fluide. Nous appellerons cet indice l'indice fluide. Une fois l'indice fluide dérivé, il est possible de développer une théorie analogue à celle de l'indice de Whittle's et nous proposons la politique de l'indice fluide comme une heuristique pour le problème originellement contraint. Cette nouvelle heuristique a plusieurs avantages sur celle de l'indice de Whittle. d'une part elle peut être dérivée explicitement et d'autre part l'indexabilité et la monotonicité du contrôle optimal (fluide) peut être aisément prouvé. Nous montrons l'applicabilité de cette heuristique en considérant plusieurs cas d'études, *e.g.*, une file d'attente multi-classe pour un serveur en présence d'abandons, l'ordonnancement opportuniste dans les réseaux sans fil, l'éco-conscience dans les fermes de serveurs, et la gestion de stock périssables. Nous observons que cette nouvelle heuristique offre de bonnes performances pour différents niveaux de charge et coincide avec l'indice de Whittle dans plusieurs cas. Dans le Chapitre 5, nous étudions le modèle d'abandon multi-classes issue du Chapitre 3, mais nous approchons dans ce cas le modèle stochastique original par un modèle de fluide. Nous menons ensuite une analyse pour le cas surchargé puis pour le cas non-surchargé. Dans le cas surchargé, nous sommes capables de caractériser pleinement la politique optimale qui n'avère être une politique de priorité stricte. En absence de surcharge, nous résolvons le problème pour deux classes en utilisant le principe du maximum de Pontryagin, puis, en se basant sur ces résultats, nous construisons une heuristique pour un nombre arbitraire de classes. La solution optimale pour un système de deux classe se trouve être une politique de seuil. Nous observons par des expérimentations numériques exhaustives que l'heuristique proposée dans ce chapitre capture correctement la structure de la politique optimale pour le problème stochastique initial. Dans la Partie III, nous approfondissons notre recherche sur le phénomène des abandons dans les systèmes de file d'attente. Contrairement aux études menées dans les parties précédentes, nous considérons maintenant une dynamique des files qui est différente de la simple structure de naissance et de mort. Plus précisément, nous autorisons la présence de départ par lots. Ceci est motivé par le problème de transmission de contenu dans le cas ou les requêtes pour un contenu peuvent être groupées pour être servies simultanément en un mode de diffusion. Nous analysons ce mode de transmission pour un serveur mono-classe de capacité infini. Dans ce chapitre nous prouvons que les politiques de seuil sont optimales en moyenne et nous dérivons la distribution de probabilité de la taille de la file d'attente pour deux configurations: (1) lorsque le taux de service est infini, et (2) lorsque le taux de service est fini. Ceci nous permet de caractériser la politique optimale de transmission (qui est de type seuil). Nous observons numériquement l'importance du calcul du seuil optimal tant les performance des politiques de seuil lorsque celui-ci n'est pas optimal sont faibles.

Laburpena euskaraz

Tesi honetan baliabideen esleipenerako problemaren kontrol dinamikoa aztertu da. Problema hauek hainbat aplikazio dituzte, adibidez, inbentarioen kudeaketa, osasun zaintza eta komunikazio sareak. Euren aplikagarritasuna dela-eta elkarrekin lehian ari diren bezero klaseei baliabideak esleitzeko problema sakon aztertua izan da literaturan. Baliabide esleipen politikak helburu ezberdinak erdiesteko diseinatu daitezke, besteak beste, baloreen kudeaketan salmentak maximizatzeko, heriotzak murrizteko osasun zerbitzuetako *triage* sistemetan edo errendimendua maximizatzeko haririk gabeko sareetan. Ikuspegi matematikotik, ilaren-teoriako ereduak baliabide esleipenerako sistemen ezaugarri asko atzeman ditzakete, eta tradizionalki ilaren sistemen errendimendua ebaluatzeko erabili izan dira, esaterako, ilaran zain dauden bezero kopuruaren banaketa aurreikusteko, itxaron denborak kalkulatzeko eta erabakiak hartzeko prozesuan laguntzeko, hau da, helburu desiragarriak lortu ahal izateko *scheduling* politikak garatzeko.

Baliabide esleipenerako problema askok zorizko izaera dute, horregatik, tesi hau kontrol estokastiko ereduetan oinarritzen da. Bereziki, *restless bandit problem*-en (RBP-en) markoan sartzen diren baliabide esleipenerako problemak aztertu dira, adibidez, uzteak gerta daitezken ilara klase-anitza, inbentario kudeaketa itemak galkorrak direnean, haririk gabeko sareetan antolamendua etab. RBP-ak *Markov Erabaki-Prozesu*-en (MEP-en) optimizazio dinamiko estokastikoaren kasu berezi bat dira. Oinarrizko problema hurrengo moduan deskribatu daiteke. Sisteman K *bandit* daude, eta gehienez horietako M aktibatu daitezke. Helburua aukeratutako errendimendu irizpidea lortzeko zein *bandit* aktibatu deskribatzen duen politika diseinatzea da. Tesi honetan, RBP-en markoan sartzen diren bi adibide multzo deskribatu dira: klaseen aukeraketa optimoa eta trafiko kargaren balantze optimoa. Bi kasu hauetan, *bandit*-ak bezero klaseen adierazle dira. Klaseen aukeraketa optimo probleman *bandit* bat aktibatzeak, *bandit* horri dagokion bezero klasea zerbitzatzea adierazten du, eta *bandit*-a pasibo uzteak, bezero klase hori ez zerbitzatzea. Ordez, trafiko kargaren balantze optimo problemetan *bandit*-a aktibatzeak, sistemara iritsi den bezero bat klase bati esleituko zaiola adierazten du, eta pasibo izateak, ez dela bezerorik esleituko klase horretara. Lehenengo kasuan helburua zein klase zerbitzatu behar den erabakitzea da eta bigarren kasuan, sistemara iritsi berri diren bezeroak zein klasetara bidali behar diren erabakitzea. Tesi honetan arreta berezia jarri da bezeroen uzteak kontsideratzen dituzten ereduetan. Bezeroen uzteak, bezeroak jasotako zerbitzuarekin ados ez daudenean edota itxarondenbora luzeegiak jasan dituztenean gertatzen dira. Fenomeno honek izugarrizko eragina du eguneroko bizitzako hainbat aplikaziotan, horren adibide dira Interneta eta telefono-dei zentroak, non bezeroak ilaran zain daudela utz dezaketen sistema edo baita zerbitzua jasotzen ari diren bitartean. Uztea oso fenomeno zitala da bai bezeroen ikuspuntutik eta baita sistemaren ikuspuntutik ere (irabazi bat galtzen baita), eta eragin ekonomikoa bat dakar. Ez da beraz harrigarria, azken urteetan ikertzaileek fenomeno hau aztertzeko izan duten irrika.

Lan honetan baliabide esleipenerako hainbat sistema aztertu dira, kasu batzuetan soluzio optimo bat lortu ahal izan da, besteetan heuristikak garatu dira hurbilketa ezberdinak jarraituz. Tesi honetan bi hurbilketa nagusi erabili dira, lehenengo hurbilketa metodo estokastikoetan oinarritzen da eta I eta III Ataletan landu da, bigarren hurbilketa kontrol optimo fluidoan oinarritzen da eta II Atalean izan da aztergai. Teknika hauek 1. Kapituluan aurkeztu dira, baita teknika hauek aplikatu diren adibideak ere.

I. Atalean RBP-en testuinguru orokorra azaldu da *bandit*-ek jaiotza-eta-heriotza motako dinamika jarraitzen duten kasurako. 2. Kapituluan indize gaitasuna eta erlaxazio Lagrangearra aurkeztu dira eta Whittle indizearen espresioa esplizitua lortu da. Espresio hau oreka-egoerako probabilitateekiko adierazten da, non probabilitate hauek jakinak diren jaiotza-eta-heriotza prozesuentzako. Whittle indizearen errepresentazio honek indize gaitasuna oreka-egoerako probabilitateen propietate gisa definitzea ahalbidetzen du. 3. Kapituluan, 2. Kapituluko emaitzak aplikatu dira uzteak gerta daitezkeen ilara baten kasurako. Politika monotonoak optimoak direla frogatu da problema baldintza-gaberako eta baita *bandit*-ak indize gaitasuna dutela. Azken honek Whittle-en indizearen espresioa lortzea ahalbidetzen du, zeina kasu partikularretan, esplizituki lor daitekeen. Bestalde, bezeroen uzteek indize gaitasunik ez duten problementzako indizeak berreskuratzen lagun dezakete, uzteak 0-rantz eramanez. Azken hau da M/M/1 ilararen kasuan indizea lortzeko erabili den teknika kostuak ganbilak diren kasurako. Whittle indizea asintotikoki optimoa dela frogatu da trafikoa geldoa eta trafikoa arina den kasuetan, eta numerikoki ikusi da Whittle indize politikaren errendimendua ona dela trafiko intentsitate ezberdinetarako. Whittle indize politika beraz optimotik gertu dagoela ondorioztatu daiteke uzteak dituen ilara baterako.

II Atalean baliabideen esleipenerako problema klase-anitza errazagoak diren problema deterministen bidez hurbildu dira. Bi hurbilketa matematiko kontsideratu dira. 4. Kapituluan I Atalean kontsideratutako erlaxazioa abiapuntutzat hartuz, problema zerbitzari-anitz klase-anitzaren bertsio erlaxatua doitu da klase bakoitzak jaiotza-eta-heriotza motako trantsizioak jasaten dituen kasurako. Doiketa honek problema determinista baldintza-gabe bat lortzea ahalbidetzen du, non W Lagrange biderkatzailea indize politika bat lortzeko erabil daitekeen, indize honi *indize fluidoa* deituko diogu. Indize hau lorturik, Whittle indizeari dagokion teoria analogoa garatu daiteke eta indize fluido politika heuristika proposatu. heuristika honen Whittle-en indize politikarekiko hainbat abantaila ditu: esplizituki lor daiteke, indize gaitasuna erraz froga daiteke, eta kontrol fluido optimoa monotonoa dela frogatzea ere erraza da. Heuristika honen aplikagarritasuna 1.2.2. Sekzioan aurkeztutako hainbat adibideen bitartez erakutsi da. Heuristika honen errendimendua ona dela erakutsi da hainbat trafiko intentsitate ezberdinetarako eta zenbait kasuetan Whittle indizearekin bat datorrela ikusi da.

5. Kapituluan, 3. Kapituluan kontsideratutako uzteak gerta daitezkeen ilara klase-anitza aztertuko da. Halere, kasu honetan jatorrizko problema estokastikoa eredu fluido batez hurbilduko da. Analisia bitan banatzen da, trafiko arinaren kasua eta trafiko geldoaren kasua. Trafikoa geldoa deneko kasuan politika optimoa guztiz karakterizatzea lortu da, politika hau lehentasun hertsiko politika bat da. Trafikoa arina deneko kasuan bi-klaseko ilararen kasua ebatzi da Pontryagin-en Maxioaren Printzipioa erabiliz eta soluzio honetan oinarrituz heuristika bat diseinatu da klase kopuru arbitrarioa den kasurako. Bi-klaseko ilararen soluzio optimoa atari motako politika bat da. Esperimentu numerikoek erakutsi dute proposatutako heuristikak jatorrizko eredu estokastikoaren propietateak ondo jasotzen dituela.

III Atalean ere uzteen fenomenoa aztertu da, baina kasu honetan ilaren dinamika ez da jaiotza-eta-heriotza motakoa baizik eta sorta irteera motako dinamikak aztertu dira. Problema hau aztertzeko motibazioa edukien banaketa problematik dator. Problema honetan eduki batentzako eskariak multzokatuak izan daitezke *multi-cast* moduan transmitituak izan daitezen. *Multi-cast* transmisio modua aztertu da uzteak gerta daitezkeen ilara klase bakar eta kapazitate infinituko zerbitzari baten kasurako. Kapitulu honetan atari politikak optimoak direla frogatu da eta ilaren luzeren banaketa kalkulatu da bi kasu ezberdinetarako: (1) zerbitzu tasa infinitua denean, eta (2) zerbitzu tasa finitua denean. Honek transmisio politika optimoa karakterizatzea ahalbidetzen du (atari motakoa dena). Numerikoki ikusi da zein den atari politika optimoa kalkulatzearen garrantzia, izan ere, optimoak ez diren atari politiken errendimendua kaxkarra da. Azkenik, A Eranskinean, soluzio bat optimoa dela frogatzea ahalbidetzen duten emaitza garrantzitsuenen bilduma aurkeztu da.

Self References

[SR1] M. Larrañaga, O.J. Boxma, R. Núñez-Queija and M.S. Squillante. “Efficient content delivery in the presence of impatient jobs”. In: *Proceedings of 27th International Teletraffic Congress* (2015).

[SR2] M. Larrañaga, U. Ayesta and I.M. Verloop. “Asymptotically optimal index policies for an abandonment queue with convex holding cost”. In: *Queueing Systems (81)2* (2015), pp.99–169.

[SR3] M. Larrañaga, U. Ayesta and I.M. Verloop. “Dynamic control of birth-and-death restless bandits: application to resource-allocation problems”. *Submitted to IEEE/ACM Transactions on Networking* (2015).

[SR4] M. Larrañaga, U. Ayesta and I.M. Verloop. “Dynamic fluid-based scheduling in a multi-class abandonment queue”. In: *Performance Evaluation (70)10* (2013), pp. 841–858, *Special Issue of IFIP Performance 2013*.

[SR5] M. Larrañaga, U. Ayesta and I.M. Verloop. “Index Policies for a multi-class queue with convex holding cost and abandonment”. In: *Proceedings of ACM SIGMETRICS* (2014), pp.125–137.

[SR6] M. Larrañaga, U. Ayesta and I.M. Verloop. “Stochastic and fluid index policies for resource allocation problems”. In: *Proceedings of IEEE Infocom* (2015), pp.1230–1238.

Bibliography

[1] S. Aalto, P. Lassila, and P. Osti. "Whittle Index Approach to Size-aware Scheduling with Time-varying Channels". In: *Proceedings of ACM Sigmetrics* (2015), pp. 57–69 (cit. on p. 9).

[2] M. Abramowitz and A.I. Stegun. *Handbook of Mathematical Functions, with Formulas, Graphs and Mathematical Tables*. Dover Publications, Inc., New York, 1965 (cit. on p. 147).

[3] D.F. Anderson. *Introduction to Stochastic Processes with Applications in the Biosciences*. University of Wisconsin at Madison, 2013 (cit. on pp. 35, 134).

[4] P.S. Ansell, K.D. Glazebrook, I. Mitrani, and J. Niño-Mora. "A semidefinite programming approach to the optimal control of a single server queueing system with imposed second moment constraints". In: *Journal of the Operational Research Society* 50 (1999), pp. 765–773 (cit. on p. 6).

[5] P.S. Ansell, K.D. Glazebrook, J. Niño-Mora, and M. O'Keeffe. "Whittle's index policy for a multi-class queueing system with convex holding costs". English. In: *Mathematical Methods of Operations Research* 57.1 (2003), pp. 21–39 (cit. on pp. 3, 6, 9, 20, 23, 26).

[6] N.T. Argon, S. Ziya, and R. Righter. "Scheduling Impatient Jobs in a Clearing System with Insights on Patient Triage in Mass-casualty Incidents". In: *Probability in the Engineering and Informational Sciences* 22.3 (2010), pp. 301–332 (cit. on p. 6).

[7] B. Ata and M.H. Tongarlak. "On scheduling a multiclass queue with abandonments under general delay costs". In: *Queueing Systems* 74.1 (2013), pp. 65–104 (cit. on p. 6).

[8] R. Atar, C. Giat, and N. Shimkin. "The $c\mu/\theta$ rule for many-server queues with abandonment". In: *Operation Research* 58.5 (2010), pp. 1427–1439 (cit. on pp. 6, 21, 39, 40, 76, 106–108, 112–114, 116).

[9] R. Atar, C. Giat, and N. Shimkin. "On the asymptotic optimality of the $c\mu/\theta$ rule under ergodic cost". In: *Queueing Systems* 67 (2011), pp. 127–144 (cit. on pp. 6, 106–108, 112–114, 116).

[10] R. Atar, A. Mandelbaum, and M.I. Reiman. "Scheduling a multi-class queue with many exponential servers: Asymptotic optimality in heavy-traffic". In: *The Annals of Applied Probability* 14 (2004), pp. 1084–1134 (cit. on p. 106).

[11] K. Avrachenkov and V. Borkar. "Whittle Index Policy for Crawling Ephemeral Content". In: *arXiv preprint arXiv: 1503.08558* (2015) (cit. on p. 9).

[12] F. Avram. "Optimal Control of Fluid Limits of Queueing Networks and Stochasticity Corrections". In: *Mathematics of Stochastic Manufacturing Systems: AMS-SIAM Summer Seminar in Applied Mathematics* 33 (1996), pp. 1–36 (cit. on p. 10).

[13] F. Avram, D. Bertsimas, and M. Ricard. “An Optimal Control Approach to Optimization of Multiclass Queueing Networks”. In: *Proceedings of Workshop on Queueing Networks of the Mathematical Inst.* (1994) (cit. on p. 10).

[14] F. Avram, D. Bertsimas, and M. Richard. “Optimization of multiclass queuing networks: a linear control approach”. In: (1994), pp. 199–234 (cit. on pp. 10, 71).

[15] U. Ayesta, P. Jacko, and V. Novak. “A nearly-optimal index rule for scheduling of users with abandonment”. In: *Proceedings of IEEE Infocom* (2011) (cit. on pp. 6, 39, 40, 107, 108, 116).

[16] F. Baccelli, P. Boyer, and G. Hebuterne. “Single-server queues with impatient customers”. In: *Advances in Applied Prabability* 16 (1984), pp. 887–905 (cit. on p. 6).

[17] N. Bäuerle. “Asymptotic optimality of tracking policies in stochastic networks”. In: *The Annals of Applied Probability* 10 (2000), pp. 1065–1083 (cit. on pp. 10, 113).

[18] N. Bäuerle and U. Rieder. “Optimal control of single-server fluid networks”. In: *Queueing Systems* 35 (2000), pp. 185–200 (cit. on p. 10).

[19] R. Bellman. *Dynamic Programming*. 1st ed. Princeton, NJ, USA: Princeton University Press, 1957 (cit. on p. 12).

[20] D.P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, 19982 (cit. on p. 9).

[21] D.P. Bertsekas. *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999 (cit. on p. 9).

[22] D.P. Bertsekas. *Dynamic programming and optimal control*. Athena Scientific, 2005 (cit. on pp. 8, 12, 13, 15, 90, 92, 158).

[23] D. Bertsimas and J. Niño-Mora. “Restless Bandits, Linear Programming Relaxations, and a Primal-Dual Index Heuristic”. In: *Operations Research* 48 (2000), pp. 80–90 (cit. on p. 3).

[24] S. Bhulai, H. Blok, and F.M. Spieksma. “K competing queues with customer abandonment: optimality of a generalized $c\mu$-rule by the smoothed rate truncation method”. In: *Work in progress* (2014) (cit. on pp. 6, 7, 36).

[25] S. Bhulai, A.C. Brooms, and F.M. Spieksma. “On structural properties of the value function for an unbounded jump Markov process with an application to a processor sharing retrial queue”. In: *Queueing Systems* 76 (2014), pp. 425–446 (cit. on pp. 12, 37, 51, 52, 132, 134, 135, 142, 143).

[26] C.F. Bispo. “The single-server scheduling problem with convex costs”. In: *Queueing Systems* 73 (2013), pp. 261–294 (cit. on pp. 20, 80).

[27] N.K. Boots and H. Tijms. “A multiserver queueing systems with impatient customers”. In: *Management Science* 45.3 (1999), pp. 444–448 (cit. on p. 6).

[28] S.C. Borst. “User-level performance of channel-aware scheduling algorithms in wireless data networks”. In: *IEEE/ACM Transactions on Networking* 13.3 (2005), pp. 636–647 (cit. on pp. 22, 80).

[29] O.J. Boxma and P.R. de Waal. “Multiserver queues with impatient customers”. In: *Proceedings of 14th International Teletraffic Congress* (1994), pp. 743–756 (cit. on p. 6).

[30] A. Brandt and M. Brandt. “On the two-class M/M/1 system under preemptive resume and impatience of the prioritized customers”. In: *Queueing Systems* 47 (2004), pp. 147–168 (cit. on p. 6).

[31] P.H. Brill and M.J.M. Posner. “Level crossings in point processes applied to queues: single-server case”. In: *Operations Research* 25.4 (1977), pp. 662–674 (cit. on p. 6).

[32] O. Brun, H. Ben Cheikh, and B. Prabhu. "Optimal speed scaling for multiclass queues". In: *HAL report: hal-01094506* (2014) (cit. on p. 10).

[33] C. Buyukkoc, P. Varaya, and J. Walrand. "The $c\mu$ rule revisited". In: *Advances in Applied Probability* 17 (1985), pp. 237–238 (cit. on pp. 20, 43, 106, 116).

[34] B.C. Chachuat. *Nonlinear and Dynamic Optimization: from theory to practice.* IC-32: Winter Semester, 2006-2007 (cit. on p. 13).

[35] M. Chaichian and M. Demichev. *Path Integrals in Physics Volume 1: Stochastic Process and Quantum Mechanics*. Taylor & Francis, 2001.

[36] J.W. Cohen. *The single server queue*. North-Holland, 1982 (cit. on p. 2).

[37] J.G. Dai. "On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models". In: *Annals of Applied Probability* 5 (1995), pp. 49–77 (cit. on p. 10).

[38] J.G. Dai and S. He. "Many-server queues with customer abandonment: A survey of diffusion and fluid approximations". In: *Journal of Systems Science and Systems Engineering* 21 (1 2012), pp. 1–36 (cit. on p. 6).

[39] J.G. Dai and D. Yao. "Brownian inventory models with convex holding cost, Part1: average-optimal controls". In: *Stochastic Systems* 3.2 (2013), pp. 442–499 (cit. on p. 6).

[40] J.G. Dai and D. Yao. "Brownian inventory models with convex holding cost, Part2: discount-optimal controls". In: *Stochastic Systems* 3.2 (2013), pp. 500–573 (cit. on p. 6).

[41] D.G. Down, G. Koole, and M.E. Lewis. "Dynamic control of a single server system with abandonments". In: *Queueing Systems* 67 (2011), pp. 63–90 (cit. on pp. 6, 35, 107).

[42] B. Doytchinov, J. Lehoczky, and S. Shreve. "Real-time queues in heavy traffic with earliest-deadline-first queue discipline". In: *The Annals of Applied Probability* 11.2 (2001), pp. 332–378 (cit. on p. 132).

[43] A.K. Erlang. "The Theory of Probabilities and Telephone Conversations". In: *Nyt Tidsskrift for Matematik B* 20 (1909) (cit. on p. 2).

[44] A. Gajrat and A. Hordijk. "Fluid approximation of a controlled multiclass tandem network". In: *Queueing Systems* 35 (2000), pp. 349–380 (cit. on pp. 10, 113).

[45] O. Garnett, A. Mandelbaum, and M. Reiman. "Designing a call center with impatient customers". In: *Manufacturing and Service Operations Management* 4 (2002), pp. 208–227 (cit. on p. 106).

[46] E. Gelenbe and I. Mitrani. *Analysis and Synthesis of Computer Systems*. London: Academic Press, 1980 (cit. on pp. 20, 43).

[47] J.M. George and J.M. Harrison. "Dynamic control of a queue with adjustable service rate". In: *Operations Research* 49.5 (2001), pp. 720–731 (cit. on p. 20).

[48] J.C. Gittins, K. Glazebrook, and R. Weber. *Multi-armed Bandit Allocation Indices*. Wiley, 2011 (cit. on pp. 3, 9, 23, 26, 27, 43, 44).

[49] K.D. Glazebrook, C. Kirkbride, and J. Ouenniche. "Index policies for the admission control and routing of impatient customers to heterogeneous service stations". In: *Operations Research* 57 (2009), pp. 975–989 (cit. on pp. 21, 23, 26, 27).

[50] K.D. Glazebrook, J. Niño-Mora, and P.S. Ansell. "Index policies for a class of discounted restless bandits". In: *Advances in Applied Probability* (2002), pp. 754–774 (cit. on p. 3).

[51] K.D. Glazebrook, P.S. Ansell, R.T. Dunn, and R.R. Lumley. "On the optimal allocation of service to impatient tasks". In: *Journal of Applied Probability* 41.1 (2004), pp. 51–72 (cit. on p. 6).

[52] K.D. Glazebrook, D.J. Hodge, C. Kirkbride, and R.J. Minty. “Stochastic scheduling: A short history of index policies and new approaches to index generation for dynamic resource allocation”. In: *Journal of Scheduling* (2013), pp. 1–19 (cit. on p. 9).

[53] C. Gromoll, P. Robert, and B. Zwart. “Fluid Limits for Processor Sharing Queues with Impatience”. In: *Mathematics of Operations Research* 33.2 (2008), pp. 375–402 (cit. on p. 6).

[54] X. Guo and O. Hernandez-Lerma. *Continuous-Time Markov Decision Processes*. Vol. 62. Springer-Verlag, 2013 (cit. on pp. 4, 11, 156).

[55] J.M. Harrison and A. Zeevi. “Dynamic scheduling of a multiclass queue in the Halfin-Whitt heavy traffic regime”. In: *Operations Research* 52.2 (2004), pp. 243–257 (cit. on p. 106).

[56] R. F. Hartl, S. P. Sethi, and R. G. Vickson. “A survey of the maximum principles for optimal control problems with state constraints”. In: *SIAM Review* 37.2 (1995), pp. 181–218 (cit. on pp. 13, 14, 157).

[57] J. Hasenbein and D. Perry (Eds). “Special Issue on Queueing Systems with Abandonments”. In: *Queueing Systems* 75.2–4 (2013), pp. 111–384 (cit. on p. 6).

[58] F. Iravani and B. Balcioğlu. “On priority queues with impatient customers”. In: *Queueing Systems* 58 (2008), pp. 239–260 (cit. on p. 6).

[59] A. Jean-Marie and E. Hyon. “Scheduling services in a queuing system with impatience and setup costs”. In: *Lecture Notes in Electrical Engineering* 62 (2010), pp. 45–50 (cit. on p. 132).

[60] B. Ji, G.R. Gupta, M Sharma, X. Lin, and N.B. Shroff. “Achieving optimal throughput and near-optimal asymptotic delay performance in multichannel wireless networks with low complexity: a practical greedy scheduling policy”. In: *IEEE/ACM Transactions on Networking* 23.3 (2015), pp. 880–893 (cit. on p. 9).

[61] D.G. Kendall. “Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain”. In: *The Annals of Mathematical Statistics* 24 (1953) (cit. on p. 2).

[62] J. Kim and A. Ward. “Dynamic scheduling of a GI/GI/1+GI queue with multiple customer classes”. In: *Queueing Systems* 75.2-4 (2013), pp. 339–384 (cit. on p. 6).

[63] L. Kleinrock. *Queueing Systems, vol. 1*. John Wiley and Sons, 1976 (cit. on p. 2).

[64] L. Kleinrock. *Queueing Systems, vol. 2*. John Wiley and Sons, 1976 (cit. on p. 2).

[65] L. Kruk, J. Lehoczky, K. Ramanan, and S. Shreve. “Heavy traffic analysis for EDF queues with reneging”. In: *The Annals of Applied Probability* 21.2 (2011), pp. 484–545 (cit. on p. 132).

[66] K. Liu, R. Weber, and Q. Zhao. “Indexability and whittle index for restless bandit problems involving reset processes”. In: *IEEE Conference on Decision and Control and European Control Conference* (2011), pp. 7690–7696 (cit. on p. 9).

[67] K. Liu and Q. Zhao. “Indexability of Restless Bandit Problems and Optimality of Whittle Index for Dynamic Multichannel Access”. In: *IEEE Transactions Information Theory* 56 (2010), pp. 5547–5567 (cit. on p. 9).

[68] C. Maglaras. “Discrete-review policies for scheduling stochastic networks: Trajectory tracking and fluid-scale asymptotic optimality”. In: *The Annals of Applied Probability* 10 (2000), pp. 897–929 (cit. on pp. 10, 113).

[69] S.P. Meyn. “Stability and optimization of queueing networks and their fluid models”. In: *Mathematics of stochastic manufacturing systems. Lectures in Applied Mathematics* 33 (1997), pp. 175–199 (cit. on pp. 10, 113).

[70] J. A. van Mieghem. “Dynamic Scheduling with Convex Delay Costs: The Generalized $c\mu$ Rule”. In: *The Annals of Applied Probability* 5.3 (1995), pp. 808–833 (cit. on pp. 6, 44, 76).

[71] J. Niño-Mora. “Restless Bandits, partial conservation laws and indexability”. In: *Advances in Applied Probability* 33.1 (2001), pp. 76–98 (cit. on p. 3).

[72] J. Niño-Mora. “Dynamic priority allocation via restless bandit marginal productivity indices”. In: *TOP* 15.2 (2007), pp. 161–198 (cit. on pp. 9, 26).

[73] J. Niño-Mora and S. Villar. “Sensor scheduling for hunting elusive hiding targets via whittle’s restless bandit index policy”. In: *5th International Conference on Network Games, Control and Optimization (NetGCooP)* (2011), pp. 1–8 (cit. on p. 9).

[74] W. Ouyang, A. Eryilmaz, and N.B. Shroff. “Asymptotically Optimal Downlink Scheduling over Markovian Fading Channels”. In: *Proceedings of IEEE INFOCOM* (2012), pp. 1–9 (cit. on p. 9).

[75] K.P. Papadaki and W.B. Powell. “Exploiting structure in adaptive dynamic programming algorithms for a stochastic batch service problem”. In: *European Journal of Operational Research* 142 (2002), pp. 108–127 (cit. on p. 132).

[76] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko. *The Mathematical Theory of Optimal Processes*. Interscience Publishers, John Wiley and Sons, Inc. New York, 1962 (cit. on p. 14).

[77] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2005 (cit. on pp. 8, 12, 55, 73, 78, 143, 156).

[78] V. Raghunathan, V. Borkar, M. Cao, and P.R. Kumar. “Index Policies for Real-Time Multicast Scheduling for Wireless Broadcast Systems”. In: *Proceedings of IEEE Infocom* (2008) (cit. on p. 9).

[79] Ph. Robert. *Stochastic Networks and Queues*. Springer-Verlag, 2003 (cit. on p. 10).

[80] S. Ross. *Introduction to stochastic dynamic programming*. Academic Press, 1983 (cit. on pp. 12, 156).

[81] A. Salch. “Ordonnancement stochastique avec impatience”. In: *PhD Thesis* (2013) (cit. on p. 7).

[82] A. Salch, J.P. Gayon, and P. Lemaire. “Optimal control of a multi-class queueing system with abandonments.” In: *MSOM, Fontainebleau* (2013) (cit. on p. 7).

[83] L.E. Schrage and L.W. Miller. “The Queue M/G/1 with the Shortest Remaining Processing Time Discipline”. In: *Operations Research* 14 (1966), pp. 670–684 (cit. on p. 20).

[84] A. Seierstad and K. Sydsæter. “Optimal control theory with economic applications”. In: *Advanced textbooks in economics* (1987) (cit. on p. 156).

[85] R. Singh, X. Guo, and P.R. Kumar. “Index policies for optimal mean-variance trade-off inter-delivery times in real-time sensor networks”. In: *Proceedings of IEEE Infocom* (2015) (cit. on p. 9).

[86] I. Taboada, F. Liberal, and P. Jacko. “An opportunistic and non-anticipating size-aware scheduling proposal for mean holding cost minimization in time-varying channels.” In: *Performance Evaluation* 79 (2014), pp. 90–103 (cit. on p. 9).

[87] I. Taboada, P. Jacko, U. Ayesta, and F. Liberal. “Opportunistic scheduling of flows with general size distribution in wireless time-varying channels”. In: *Proceedings of 26th International Teletraffic Congress* (2014) (cit. on p. 9).

[88] H. Takagi. *Queueing analysis: a foundation of performance evaluation*. Queueing Analysis: A Foundation of Performance Evaluation vol. 2. North-Holland, 1993 (cit. on p. 2).

[89] D. Towsley and S.S. Panwar. “On the optimality of Minimum Laxity and Earliest Deadline First scheduling for real-time multiprocessors”. In: *Proceedings of IEEE EUROMICRO-90 Real Time Workshop* (1990), pp. 17–24 (cit. on p. 132).

[90] E. Trélat. “Contrôle optimal: théorie et applications”. In: *Vuibert, Collection : Mathématiques Concrètes* (2008) (cit. on p. 13).

[91] M.H. Veatch and L.M. Wein. “Scheduling a make-to-stock queue: index policies and hedging points”. In: *Operations Research* 44.4 (1996), pp. 634–647 (cit. on pp. 7, 23, 87).

[92] M.H. Veatch and L.M. Wein. “Optimal dynamic scheduling policy for a make-to-stock production system”. In: *Operations Research* 45.1 (1997), pp. 42–53 (cit. on p. 7).

[93] I.M. Verloop. “Asymptotically optimal priority policies for indexable and non-indexable restless bandits”. In: *HAL Report, hal-00743781* (2015) (cit. on pp. 9, 45).

[94] I.M. Verloop and R. Núñez-Queija. “Asymptotically optimal parallel resource assignment with interference”. In: *Queueing Systems* 65.1 (2010), pp. 43–92 (cit. on pp. 10, 113).

[95] R.R. Weber and S.Jr. Stidham. “Optimal Control of Service Rates in Networks of Queues”. In: *Advances in Applied Probability* 19.1 (1987), pp. 202–2018 (cit. on p. 6).

[96] R.R. Weber and G. Weiss. “On an index policy for restless bandits”. In: *Journal of Applied Probability* 27 (1990), pp. 637–648 (cit. on pp. 3, 9, 45).

[97] G. Weiss. “On optimal draining of reentrant fluid lines”. In: *Stochastic Networks, eds. F.P. Kelly and R.J. Williams* (1995), pp. 91–103 (cit. on pp. 10, 71).

[98] W. Whitt. “Efficiency-Driven Heavy-Traffic Approximations for Many-Server Queues with Abandonments”. In: *Management Science* 50 (2004), pp. 1449–1461 (cit. on p. 6).

[99] P. Whittle. “Restless Bandits: Activity Allocation in a Changing World”. In: *Journal of Applied Probability* 25 (1988), pp. 287–298 (cit. on pp. 3, 8, 9, 22, 24, 27).

[100] A. Wierman, L.L.H. Andrew, and A. Tang. “Power-aware speed scaling in processor sharing systems”. In: *Proceedings of IEEE Infocom* (2009) (cit. on pp. 21, 83).

Baliabideen esleipenerako sistema estokastiko eta fluidoen kontrol dinamikoa

Maialen LARRAÑAGA

Laburpena

Tesi honetan hainbat alorretan agertzen diren baliabideen banaketarako sistemen kontrola aztertu da. Aplikazio alor gisa, besteak beste, inbentario kudeaketa, osasun zaintza eta komunikazio sareak aipa daitezke. Oinarrizko helburua, elkarren aurka lehian ari diren proiektu edo entitateen artean baliabideak efikaziaz banatzea da. Mota honetako problemak natura estokastikoa dute eta oso konplexuak gerta daitezke ebazterako orduan. Hori dela eta, tesi hau errendimendu ona erakusten duten heuristikak lortzean oinarritzen da. Lehenengo atalean, *Restless Bandit Problem*-en markoa kontsideratu dugu. Restless Bandit Problem-ak optimizatio problema dinamiko estokastikoen kasu partikular bezala ikus daitezke. Lagin-ibilbidearen gaineko baldintzak erlaxatuz indizeetan oinarritutako heuristikak definitu daitezke problema baldintzatu originalarentzako. Heuristika hau Whittle indize politika izenez ezagutzen da. Tesi honetan, proiektu edo entitateak jaiotza-eta-heriotza motako prozesu bidez deskriba daitezkeneen kasurako, Whittle indizea ebatzi da esplizituki. Adierazpen esplizitu hau ebatzi ahal izateko hainbat baldintza tekniko egiaztatu behar dira. Ez hori bakarrik, kasu partikularretan bakarrik garatu daiteke. Bestalde, uzteak gerta daitezken ilara klase-anitz bat aztertu da. Problema honentzako Whittle indeze politika garatu da eta asintotikoki optimoa dela frogatu da trafikoa arina edo geldoa den kasuetan. Bigarren atalean, baliabideen banaketarako problema estokastikoa, determinista den problema fluido batekin hurbildu da. Lehenik eta behin, lehenengo atalean azaldutako optimizazio problema erlaxatuaren bertsio fluidoa aurkeztu da, ondoren indize fluido politika garatu da. Indize fluidoa esplizituki kalkula daiteke kasu guztietan eta beraz, Whittle indizea ebazterakoan topatutako arazo teknikoak gainditzen ditu. Bai Whittle indize politika eta bai indize fluido politika hainbat sistementzako ebatzi dira, adibidez, kontzientzia-energetikoz hornitutako zerbitzari-parkeak, antolamendu oportunista haririk gabeko sistemetan, produktu galkorren inbentarioen planifikazioan etab. Numerikoki erakutsi da bi indize politikak optimotik hurbil daudela. Bigarren, antolamendu kontrol optimoa aztertu da uzteak gerta daitezkeen ilara klase-anitz baterako. Kontrol fluido optimoa garatu da bi bezero klasek baliabide bakarrarengatik elkar lehiatzen duten kasurako. Bi klaseko kasutik lortutako kontrol optimoan oinarrituz klase-anitzeko ilarentzako heuristika bat proposatu da. Heuristika honen errendimendua, trafikoa geldoa denean, sistema estokastiko originalari dagokion optimotik gertu dagoela ikusi da. Azkenik, hirugarren atalean, bezeroen abandonua aztertu da edukien esleipen problema baterako. Eduki berdina eskatu duten bezeroak taldekatu egin daitezke, edukia efizienteki igorria izan dadin *multi-cast* moduan. Problema hau optimoki ebatzi da, eta politika optimoa *threshold* edo *atari* motakoa dela ikusi da.

Gaien Aurkibidea

1 Sarrera **1**

1.1 Ilaren teoriaren aipamen laburra 2

1.2 Eredu matematikoak 3

1.2.1 Restless bandit problem-a 3

1.2.2 Klase-anitzeko baliabide-esleipen problemen adibideak 4

1.2.3 Bezeroen uzteak 6

1.3 Metodologia 8

1.3.1 Erlaxazio Lagrangearra 8

1.3.2 Fluidoen bidezko hurbilketa 10

1.3.3 Kontrol estokastiko optimoa 10

1.3.4 Kontrol optimo determinista 13

1.4 Tesiaren laburpena 15

I Baliabideen esleipenerako sistema estokastikoen kontrol dinamikoa **17**

2 Indize politika jaiotza-eta-heriotza motako Restless Bandit-entzat **19**

2.1 Sarrera 19

2.2 Eredu matematikoa 20

2.2.1 Adibideak 21

2.3 Erlaxazioa eta indize gaitasuna 22

2.3.1 Atari-politikak 23

2.3.2 Indize-gaitasuna eta Whittle indizea 24

2.4 Whittle indize politika 27

2.5 Whittle indize politikaren errendimendua 27

2.6 Eranskina 28

2.6.1 2.1. Proposizioaren frogapena 28

2.6.2 2.2. Proposizioaren frogapena 28

2.6.3 2.1. Korolarioaren frogapena 31

3 Indize politika uzteak gerta daitezken ilara batentzat **33**
3.1 Ereduaren deskribapena eta aurrekariak . . . 34
3.2 Erlaxazioa . . . 36
3.2.1 Atari-politikak . . . 36
3.2.2 Indize-gaitasuna . . . 38
3.2.3 Whittle indizea . . . 38
3.3 Kasu bereziak . . . 39
3.3.1 Mantentze-kostu linealak . . . 39
3.3.2 Mantentze kostu ganbilak . . . 41
3.4 M/M/1 ilara klase-anitza . . . 43
3.5 Errendimendu optimoa asintotikoki . . . 45
3.5.1 Zerbitzari anitzeko sistema . . . 45
3.5.2 Trafiko arina eta trafiko geldoa . . . 46
3.6 Emaitza numerikoak . . . 46
3.6.1 Politika ezberdinen egitura . . . 47
3.6.2 Errendimenduaren ebaluazioa . . . 48
3.7 Eranskina . . . 52
3.7.1 3.1. Proposizioaren frogapena . . . 52
3.7.2 3.3. Proposizioaren frogapena . . . 56
3.7.3 3.4. Proposizioaren frogapena . . . 57
3.7.4 3.5. Proposizioaren frogapena . . . 60
3.7.5 3.6. Proposizioaren frogapena . . . 62
3.7.6 3.7. Proposizioaren frogapena . . . 64
3.7.7 3.8. Proposizioaren frogapena . . . 65
3.7.8 3.9. Proposizioaren frogapena . . . 67

II Baliabideen esleipenerako sistema fluidoen kontrol dinamikoa **69**

4 Indize fluido politika jaiotza-eta-heriotza motako *restless bandit*-entzat **71**
4.1 Optimizazio problema erlaxatuaren bertsio fluidoa . . . 72
4.1.1 Eredu fluidoa eta *bias*-optimoa izatea . . . 72
4.1.2 Kontrol optimo fluidoa eta indize fluidoa . . . 74
4.1.3 Indize fluido politika . . . 76
4.2 Indize estokastikoa eta fluidoaren arteko baliokidetasuna trafiko arinean . . . 76
4.3 Adibideak . . . 78
4.3.1 Uzteak gerta daitezkeen ilara klase-anitzean *scheduling*-a . . . 78
4.3.2 Haririk gabeko sare batean *scheduling* oportunista . . . 80
4.3.3 Blokeatzea/bideratzea kontzientzia energetikoz hornitutako zerbitzari-parketan . . . 83
4.3.4 Inbentarioen kudeaketa itemak galkorrak direnean . . . 87
4.4 Eranskina . . . 90
4.4.1 4.1. Proposizioaren frogapena . . . 90
4.4.2 4.2. Lema . . . 96

4.4.3 4.2. Proposizioaren frogapena 99
4.4.4 4.3. Proposizioaren frogapena 103
4.4.5 4.4. Proposizioaren frogapena 103

5 Kontrol fluido optimoa uzteak gerta daitezken ilara batean 105
5.1 Sarrera 106
5.2 Ereduaren deskribapena 107
5.3 Kontrol fluido eredua 108
5.3.1 Lan karga txikian politika optimoa bi bezero klaserako 109
5.3.2 Lan karga handia denean politika optimoa bezero klase kopuru arbitrariorako 111
5.3.3 Kontrol optimoa eredu estokastiko eta fluidoaren arteko alderatzea 112
5.4 Bezero klase kopuru arbitrarioa 114
5.5 Emaitza numerikoak 116
5.5.1 Errendimenduaren ebaluatzea bi bezero klaserentzat 116
5.5.2 Errendimenduaren analisia bezero klase kopuru arbitrariorako 119
5.6 Eranskina 120
5.6.1 5.1. Lemaren frogapena 120
5.6.2 5.1. Proposizioaren frogapena 122

III Sistema estokastikoen kontrol dinamiko optimoa 129

6 Uzteak gerta daitezkeen sorta-ilarak 131
6.1 Sarrera 132
6.2 Ereduaren deskribapena 133
6.3 Atari-politika optimoak 135
6.4 Banaketa egoera-egonkorrean 135
6.4.1 Zerbitzu tasa infinitua 136
6.4.2 Zerbitzu tasa finitua 136
6.5 Atari-politika optimoaren karakterizazioa 138
6.6 Adibideak 139
6.7 Eranskina 142
6.7.1 6.1. Proposizioaren frogapena: $\mu < \infty$ kasua 142
6.7.2 6.2. Proposizioaren frogapena 146
6.7.3 6.3. Proposizioaren frogapena 150
6.7.4 6.1. Lemaren frogapena 151
6.7.5 6.2. Lemaren frogapena 153

A Eranskina 155
A.1 Politika optimo egonkorren existentziareko baldintzak: Bellman ekuazioa 155
A.2 Beharrezko baldintzak: Pontryagin-en Printzipio minimioa 156
A.3 Baldintza nahikoak: Hamilton-Jacobi-Bellman Ekuazioa 157

Autorearen artikuluak 159

Bibliografia **161**

1. Kapitulua

Sarrera

Contents

1.1 Ilaren teoriaren aipamen laburra **2**
1.2 Eredu matematikoak **3**
1.3 Metodologia **8**
1.4 Tesiaren laburpena **15**

Elkarrekin lehian ari diren bezero klaseei baliabideak esleitzeko problema sakon aztertua izan da literaturan, dituen hainbat aplikazio direla eta, adibidez, inbentarioen kudeaketa, osasun zaintza eta komunikazio sareak. Baliabide esleipen politikak helburu ezberdinak erdiesteko diseinatu daitezke, besteak beste, baloreen kudeaketan salmentak maximizatzeko, heriotzak murrizteko osasun zerbitzuetako *triage* sistemetan edo errendimendua maximizatzeko haririk gabeko sareetan. Ikuspegi matematikotik, ilaren-teoriako modeloak baliabide esleipenerako sistemen ezaugarri asko atzeman ditzakete, eta tradizionalki ilaren sistemen errendimendua ebaluatzeko erabili izan dira, esaterako, ilaran zain dauden bezero kopuruaren distribuzioa aurreikusteko, itxaron denborak kalkulatzeko eta erabakiak hartzeko prozesuan laguntzeko, hau da, helburu desiragarriak lortu ahal izateko antolamendu politikak garatzeko.

Baliabide esleipenerako problema askok zorizko izaera dute, horregatik, tesi honetan kontrol estokastiko ereduetan oinarritu gara, hau da, *Markov-en Erabaki Prozesuak Denbora Jarraian* (MEPDJ) ereduetan. Kasu batzuetan kontrol optimoa zehaztu daiteke. Hala ere, kasu gehienetan ezin da soluzio optimorik aurkitu eta errendimendu ona erakusten duten heuristikak garatzea da helburua. Lan honetan baliabide esleipenerako hainbat sistema aztertu dira, kasu batzuetan soluzio optimo bat lortu ahal izan da, besteetan heuristikak garatu dira.

1.1. Sekzioan ilaren teoriaren aipamen labur bat aurkeztu da. 1.2. Sekzioan tesian erabili diren ereduak motibatu dira. Ondoren, 1.3. Sekzioan, tesian jorratu diren metodoak aurkeztu dira eta azkenik, 1.4. Sekzioan tesiko gainerako kapituluen laburpen bat azaldu da.

1.1 Ilaren teoriaren aipamen laburra

Ilaren teoriako ereduak telekomunikazioen arloan sortu ziren, A.K. Erlang matematikari danimarkarrak telefono-zentralak irudikatzeko garatu zituenean [42]. Ilara ereduak oso erabiliak izan dira Ikerketa Operatibo arloko problema anitz aztertzeko. Gaur egungo Ikerketa Operatiboa bigarren gerrate handian sortu zen, Erresuma Batuko gobernuak eremu hau erabili zuenean hainbat arazo militarrei aurre egiteko. Beranduago, garraio sistemetara, fabrikazio industriara eta osasun zaintzara zabaldu zen. 50. urte bukaeran, lehenengo ordenagailuen sorrerarekin batera, ilaren ereduak ordenagailuen errendimendua ebaluatzeko erabiltzen hasi ziren, erabakiak hartzeko problema konplexuetan lagun zezaten. Urte askotan Ikerketa Operatibo Estokastikoa izugarri arrakastatsua bilakatu da informatikari eta matematikari aplikatuen artean hainbat problema ebazteko, adibidez, haririk gabeko sareetan sortutako errendimendu arazoei aurre egiteko, sare sozialak ulertzeko, informazio transferentzia arintzeko etab. Ilaren teoriaren influentziaren erakusgarri da *Queueing Systems* aldizkaria, 1986. urtean sortu zen ilara ereduak dituzten aplikazioak jorratzeko.

Ilara ereduak, sistema errealistak irudikatzeko erabili badira ere, nahiko errazak dira eta analisi matematiko bat garatu daiteke. Honek, sistema ezberdinen errendimendua ulertu eta ebaluatzea ahalbidetzen du. Ilaren teoriari buruzko informazio gehiago aurki daiteke honako lanetan [62, 63, 87].

Ilaren eredu bat hurrengo elementuek definitzen dute: sarrera prozesu bat, irteera prozesu bat, eta zerbitzu diziplina bat. Sarrera prozesuak, bezeroen iritsiera deskribatzen du, irteera prozesuak bezero bat zerbitzatzeko ematen den denbora deskribatzen du, eta zerbitzu diziplinak ilaran zain dauden bezeroak zein hurrenkeratan zerbitzatu erabakitzen du. Prozesu hauek zorizko izaera izan dezakete. D. G. Kendall [60] matematikariak hurrengo notazioa erabili zuen ilara ereduak deskribatzeko:

$$A/B/s/C,$$

non A-k iritsiera-arteko denboren distribuzioa adierazten duen, B-k zerbitzu denboraren distribuzioa, s-k sisteman erabilgarriak diren zerbitzari kopurua adierazten du eta C-k sistemaren kapazitate osoa, hau da, sistemak jaso ditzakeen bezeroen kopuru maximoa. Une honetan ohar semantiko bat beharrezkoa da, izan ere, *bezero* hitza tesian zehar erabiliko da hainbat gauza adierazteko, adibidez, item baten eskaria, egingizunak, paziente edo erabiltzaileak, eta *zerbitzari* hitza prozesatze ahalmena duten entitateak adierazteko, adibidez, sendagileak, operadoreak edo makinak.

Ilarak tipikoki Markov kateen bidez irudikatu izan dira. Markov kateak zorizko gertakizunen sekuentziak dira eta sistemaren une horretako egoerarekiko soilik dute menpekotasuna. Propietate hau *memoryless* izenez da ezaguna. Markov kateak oso erabiliak izan dira hainbat alorretan sistema fisiko edo biologikoen adierazpen gisa, adibidez, populazio prozesuen ereduak egiteko edo ahots errekonozimendurako. Markov kateen adibide kanonikoa M/M/1 ilara da, non M letra iritsiera-arteko denbora eta zerbitzu denboraren distribuzio esponentzialaren adierazgarri den eta 1 zenbakiak zerbitzari bakarra adierazten duen. Eredu hau sakon aztertua izan da eta tesi honetan rol garrantzitsu bat jokatuko du. Zerbitzari bakarreko ilaren inguruko informazio gehiago lortzeko ikusi [35].

1.2 Eredu matematikoak

Tesi honetan baliabideen esleipenerako kontrol dinamikoa aztertu da. *Restless bandit problem*-en (RBP-en) markoa kontsideratu da, non RBP-a baliabide esleipenerako problemen kasu orokor bat den, ikusi 1.2.1. Sekzioa. Bereziki, bi motatako baliabide esleipenerako problema klase-anitz aztertu dira, hau da, alde batetik klaseen aukeraketa optimoa eta bestetik trafikoaren kargaren balantze optimoa. Azken hauek 1.2.2. Sekzioan aurki daitezke. Tesi honetan arreta berezi bat jarri da bezeroen abandonuak kontsideratzen dituzten baliabide esleipen problemetan. Hori dela eta, 1.2.3. Sekzioan abandonuaren inguruko aipamen bat eta abandonuak kontsideratzen dituzten ereduak aurkeztu dira.

1.2.1 Restless bandit problem-a

RBP-ak *Markov Erabaki-Prozesu*-en (MEP-en) optimizazio dinamiko estokastikoaren kasu berezi bat dira. Oinarrizko problema hurrengo moduan deskribatu daiteke. Sisteman K *bandit* daude, eta gehienez horietako M aktibatu daitezke. Helburua aukeratutako errendimendu irizpidea lortzeko zein bandit aktibatu deskribatzen duen politika diseinatzea da. Hurrengo elementuek definitzen dute k *bandit*-a: egoera espazio bat $E_k = \{0, 1, \ldots\}$, akzio espazio bat $A_k = \{0, 1\}$, non 0 akzioak k bandit-a pasiboa dela adierazten duen eta 1 akzioak k bandit-a aktiboa dela, trantsizioa tasak $q_k^a(m, \tilde{m})$ non $m, \tilde{m} \in E_k$ eta $a \in \mathcal{A}_k$, eta kostu funtzioa $C_k(m, a)$ non $k = 0, 1, \ldots K$. Azken hau sistemaren egoeraren $m \in E_k$ eta akzioaren $a \in \mathcal{A}_k$ menpekoa da. Bandit-ak elkarrekiko dependentzia dute, denbora une batean gehienez M bandit aktibatu baititzake kontrol politikak.

Sistemaren egoera honela definitzen da, $\vec{N}^\phi(t) := (N_1^\phi(t), \ldots, N_K^\phi(t))$, non $N_k^\phi(t) \in E_k$ k bandit-aren egoeraren adierazpide den ϕ politika kontsideratuko balitz.

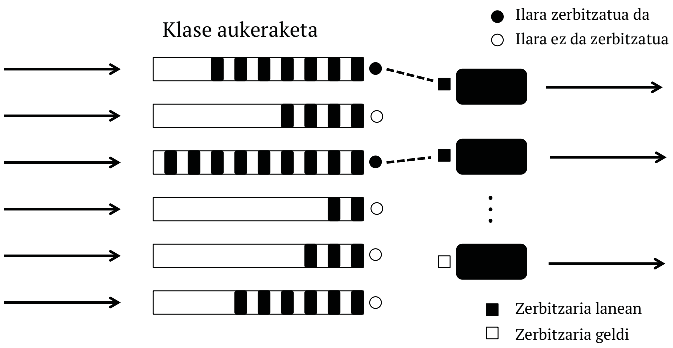
Demagun $S_k^\phi(\vec{N}^\phi(t)) \in \{0, 1\}$ aldagaiak k bandit-arekiko hartu den erabakia adierazten duela sistemaren egoera $\vec{N}^\phi(t)$ denean. Orduan, helburua epe luzerako batez besteko itxaropena minimizatzen duen ϕ esleipen politika aurkitzea da, hau da,

$$\min_{\phi} \limsup_{T\to\infty} \frac{1}{T} \mathbb{E} \left(\int_0^T \sum_{k=1}^K C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t)))\mathrm{d}t \right). \tag{1.2.1}$$

Ez hori bakarrik, erabilgarriak diren baliabideen gaineko hurrengo baldintza denbora unitate guztietan beteko da

$$\sum_{m=1}^K S_k^\phi(\vec{N}^\phi(t)) \leq M. \tag{1.2.2}$$

P. Whittle ikertzailea izan zen RBP-en markoa aurkeztu zuen lehena [98] artikuluan, *Multi-Armed Bandit Problem*-en (MABP-en) orokortze gisa. MABP-a batean $M = 1$ dela onartzen da, hau da, denbora unitate bakoitzean sistemaren planifikatzaileak bandit bat aktibatuko du. Aktibatua izan den bandit-ak kostu bat dakarkio sistemari eta bere egoera estokastikoki aldatzen da. Aktibatuak izan ez diren bandit-en egoera *izoztuta* mantentzen da. Gittins ikerlariak soluzio berritzaile batean frogatu zuen MABP-a ebazten duen soluzio optimoa indize-politika bat dela, gaur egun Gittins indize-politikatzat ezaguna dena [47]. Indize-politika edo erregelak k bandit-ari $G_k(m_k)$ funtzioa esleitzen dio non $k \in \{1, \ldots, K\}$ eta $G_k(\cdot)$-k k bandit-aren parametroekiko soilik duen dependentzia. Orduan indize-politika optimoak $\vec{m}$ egoeran



Irudia 1.1: Bezero klaseei baliabideak esleitzearen problema

$G_k(m_k)$-rik handiena duten bandit-ak aktibatzea erabakitzen du. Interes praktikoa duten problema asko ezin sar daitezke MABP-ren markoan, izan ere, k bandit-aren egoera aldakorra izan daiteke nahiz aktibatua ez izan eta kostu bat dakarkio sistemari. Kasu honetan sistema RBP markoan sartzen da. Orokorrean, RBP-en arloan soluzio optimoa karakterizatzea ez da posible.

RBP-en teoriak izugarrizko bultzada jaso du azken urteetan K. Glazebrook, J. Nino-Mora eta R. Weber bezalako ikertzaileei esker, ikusi adibidez [5, 23, 47, 49, 71, 95].

Tesi hau bandit-en egoera aldaketek dinamika berezi bat jasaten duen RBP-en kasu batean oinarritzen da. Bi prozesu mota kontsideratu dira, jaiotza-eta-heriotza prozesuak, hau da, bandit bakoitzaren egoerak bi aldaketa jasan ditzake, *jaiotzak*, non $q_k^a(m, m+1) > 0$ eta $m \in E_k$, edo *heriotzak*, non $q_k^a(m, m-1) > 0$ eta $m \in \{1, 2, \ldots\}$. Bigarren prozesuak bandit bakoitzaren egoerak hiru aldaketa izatea ahalbideratzen du, hau da, *jaiotzak*, *heriotzak* edo m tamainako *sorta irteerak*, hau da, $q_k^a(m, 0) > 0$, eta $m \in E_k$. Azken hau 6. Kapituluan kontsideratu da, non zerbitzariaren kapazitate infinitua onartu dugun eta zerbitzaria aktibatzen denean ilaran zain dauden bezero guztiak une berean zerbitzua jasotzen duten.

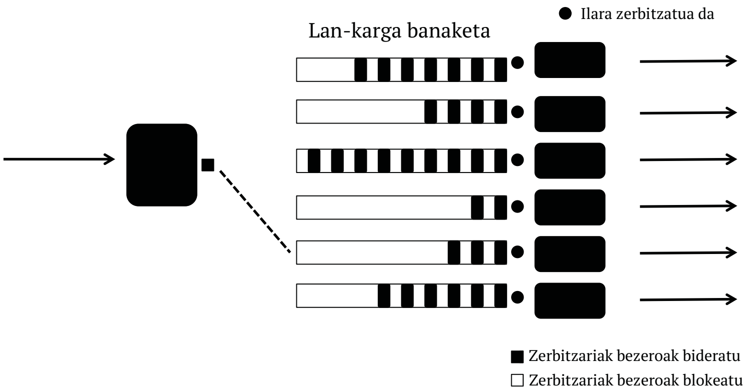
Tesi honetan aztertu ditugun sistemak egonkorrak dira, beraz, epe luzerako batez besteko kostuaren itxaropena, orekan sistemak duen batazbesteko kostuarekin bat dator. Sistema bat egonkorra izan dadin bete behar dituen baldintzak [53]-en aurki daitezke.

1.2.2 Klase-anitzeko baliabide-esleipen problemen adibideak

Sekzio honetan, 1.2.1. Sekzioan aurkeztutako RBP-en markoan sartzen diren eta tesi honetan garrantzi bezeria duten bi adibide multzo deskribatuko dira lehenik eta behin. Hau da, klaseen aukeraketa optimoa eta trafiko kargaren balantze optimoa. Bigarren, tesi honetan kontsideratu diren kostu funtzioen eztabaida labur bat aurkeztu da.

Klaseen aukeraketa optimoa

Klaseen aukeraketa optimo problemetan, klase k-ko bezeroak klase horri dagokion ilarara iristen dira. Kasu honetan erabilgarriak diren baliabideak zerbitzariak dira, eta M dira guztira. Helburua M zerbitzari horiek elkar lehian ari diren K bezero klaseei esleitzea da. Zerbitzari bakoitza klase bakarrari esleitu dakioke, eta klase bakoitzak gehienez zerbitzari batengandik jaso dezake zerbitzua. k klaseko bezero bat sistematik irteten da banaketa esponentzialeko denbora tarte baten ostean, non trantsizio edo aldaketa tasa zerbitzaria k klaseari esleitu zaion ala ez eta k bandit-aren egoeraren (*i.e.,* bezero kopuruaren) araberakoa



Irudia 1.2: Trafiko-karga orekatzearen problema

den. Zerbitzariak klaseei esleitzeko erabakia bezeroak ilaran mantentzeak dakarren kostuaren araberakoa eta beste kostu estraren araberakoa izango da. Klaseen aukeraketa optimoaren adibide gisa ikusi 1.1. Irudia. Problema hau RBP-en kasu berezi bat da: bandit batek bezeroen klase bat adierazten du eta k bandit-aren egoera k klasean zain dauden bezero kopurua. Bestalde, problema hauetan bandit-en egoeraren aldaketa jaiotza-eta-heriotza motakoa da. Testuinguru honetan $S_k^{\phi}(\vec{N}^{\phi}(t))$ akzio aldagaiak zein k bandit aktibatuko den erabakitzen du, hau da, $S_k^{\phi}(\vec{N}^{\phi}(t)) = 1$ baldin k klasea zerbitzatua bada eta $S_k^{\phi}(\vec{N}^{\phi}(t)) = 0$ baldin k klaseak zerbitzurik ez badu jasotzen. 2.2.1. Sekzioan modelo hau sakonago aztertu da.

Marko hau oso erabilia da adibidez telefono-dei zentroak irudikatzeko. Telefono-dei zentroetan klase bateko bezeroak, non klasea bezeroak nahi duen informazio edo laguntza adieraz dezakeen, itxaronaldi-moduan jartzen dira operadore bat deia erantzuteko gai den arte. Zentro batek eskura dituen operadoreen kopurua finitua denez, klase ezberdinetako bezeroak ilaran itxaron behar dute. Zentroak erabaki bat hartu behar du, zein bezero klaseri zerbitzua eman aukeratuz zentroaren irabaziak edo bezeroen esperientzia maximizatzeko.

Multzo honen barruan erortzen diren hiru adibide kontsideratu dira tesi honetan: ilarra zerbitzari-bakar klase-anitza bezeroen abandonuekin (ikusi 3. Kapitulua, 4.3.1. Sekzioa eta 5. Kapitulua), haririk gabeko *downlink* katea planifikazio oportunistarekin (ikusi 4.3.2. Sekzioa) eta *multi-cast* transmisio ilara bat bezeroen uzteekin (ikusi 6. Kapitulua).

Trafikoaren karga banaketa

Trafiko-karga banaketa problemetan, bezeroak sistemara iritsi orduko bidaltzaile batek bezeroa K ilaratako bati esleitzen dio. Bidaltzaile honek erabakitzen du zein K ilaratara joango den iritsi berri den bezeroa, ala bezeroa sistematik blokeatzea. Kasu honetan bidaltzailea litzateke errekurtsoa eta beraz problema mota honetan $M = 1$. Bezero batek, K ilaretako batean sartu denean, distribuzio esponentzialeko denbora tarte bat behar du sistematik irteteko. Irteera tasa ilara horretan dauden bezero kopuruaren araberakoa da. Hartu beharreko erabakia beraz, iritsi berri den bezeroa sisteman onartua den ala ez da, eta onartua bada zein ilaratara bidali behar den. Problema hau RBP-en testuinguruan sartzen da, non bandit batek bezero kopurua adierazten duen, eta k bandit-aren egoerak zenbat k klaseko bezero dauden adierazten du. Bandit baten egoeraren aldaketa jaiotza-eta-heriotza motakoa da. Sistemara iritsi den bezero bat k klaseko ilarara bidali bada k bandit-a aktibatu dela esango da, hau da, existitzen dela $k \in \{1, \ldots, K\}$

non $S_k^{\phi}(\vec{N}^{\phi}(t)) = 1$ eta $S_j^{\phi}(\vec{N}^{\phi}(t)) = 0$, $j \neq k$, edo sistematik blokeatu bada, hau da, $S_k^{\phi}(\vec{N}^{\phi}(t)) = 0$ eta $k \in \{1, \dots, K\}$, orduan bandit-ik ez dela aktibatu esango da. Mekanismo honek ilaretan dagoen trafiko karga orekatzen laguntzen du, ikusi 1.2. Irudia adibide gisa. 2.2.1. Sekzioan problema hau sakonki azaldu da.

Trafiko-karga banaketaren adibide klasiko bat zerbitzari-parkeena da. K zerbitzari ezberdin ari dira lanean bezeroei zerbitzua eskainiz, bezeroak sistemara iristean bidaltzaileak zerbitzari bati dagokion ilarara bidaltzen ditu. Bezeroa lanean ari den zerbitzari batera bidali bada orduan ilaran itxaron beharko du zerbitzaria berarekin lanean hasteko gai den arte.

Klase-aukeraketa problemetan ez bezala, trafiko-karga banaketa problemetan erabakia bidaltzaileak hartzen du eta ez zerbitzariak. Tesi honetan bi eredu aztertuko ditugu trafiko-karga banaketa problema gisa deskriba daitezkeenak: energia-kontzientzia duen zerbitzari-parke baten antolamendua, ikusi 4.3.3. Sekzioa, eta inbentario kudeaketa problema bat item galkorrekin, ikusi 4.3.4. Sekzioa.

Kostu-funtzioaren gaineko hipotesiak

1.2.1. Sekzioan tesi honetan kontsideratu den eredu orokorra aurkeztu da sekzio honetan bada kostu-funtzioaren gaineko hipotesiak zehaztu dira.

$C_k(m, a)$ kostu-funtzioa, non $k \in \{1, \dots, K\}$, $m \in E_k$ eta $a \in \{0, 1\}$ modeloaren araberakoa da, orokorrean halere, mantentze kostua eta beste kostu estren arteko batura adierazko du. Kostu estrak eredu bakoitzean kontsideratu diren fenomeno ezberdinen araberakoak izango dira. Fenomeno horiek sortutako kostuak zera dira: (i) bezeroak zerbitzua jaso aurretik sistema uzten duenean sortutako kostua, (ii) energia kontsumoak ekarritako kostua, honen adibide zerbitzari-parkeetan abiadura-azkartze politikak erabiltzea aipa daiteke, kasu honetan zerbitzariak zenbat eta bezero gehiago izan ilaran itxaroten orduan eta azkarrago egiten du lan, (iii) bezeroak blokeatzean sortutako kostua, bezeroei sisteman sartzea debekatzen zaienean sortzen da kostu hau, eta (iv) zerbitzaria pizte kostua, sistema batzuetan zerbitzariak martxan jartzeak kostu estra dakar. Azken lau fenomeno hauek tesian aztertuak izan dira. Bestalde, $C_k(m, a)$ kostua m-n konbexua dela onartu dugu non $a = 0, 1$, eta $k \in \{1, \dots, K\}$. Kostuak konbexuak izatearen hipotesia ilaren teoriako sistema sarritan kontsideratu da, adibidez [5, 38, 39, 94]. Van Mieghem-ek, [69]-en, kostuak linealak izateak dakartzan desabantailak eztabaidatu zituen. Kostuak konbexuak direla onartuz merkatuaren ospea edo bezeroen sinesgarritasuna irudika daitezke besteak beste. [4]-en autoreek kostu linealak kontsideratzearen desabantailak erakutsi zituzten zerbitzari-bakar klase-biko ilara batean, lehentasun txikiena duen ilara ikaragarri hazten baita.

1.2.3 Bezeroen uzteak

Tesi honetan arreta berezia jarri da bezeroen uzteak kontsideratzen dituzten ereduetan. Uzteak, bezeroak jasotako zerbitzuarekin ados ez daudenean edota itxarondenbora luzeegiak jasan dituztenean gertatzen dira. Fenomeno honek izugarrizko inpaktua du eguneroko bizitzako hainbat aplikaziotan, horren adibide dira Interneta eta telefono-dei zentroak, non bezeroak ilaran zain daudela sistema utzi dezaketen edo baita zerbitzua jasotzen ari diren bitartean. Uzteak oso fenomeno zitalak dira bai bezeroen ikuspuntutik eta baita sistemaren ikuspuntutik ere (irabazi bat galtzen baita), eta inpaktu ekonomikoa bat dakar. Ez da beraz harrigarria, azken urteetan ikertzaileek uzteen fenomenoa aztertzeko izan duten irrika. Honen erakusle da *Queueing Systems* aldizkariaren argitalpen berezia, non uzteak jorratzen dituzten ilara-ereduak aztertzen diren [56], baita [37] ikerketa-artikulua ere, zerbitzari-aniztun ilarak aztertzen dituena. Bereziki garrantzitsuak dira bezeroen uzteak ilara-ereduetan duen inpaktua eta errendimendua aztertzen duten ikerketa-artikuluak, ikusi [7, 16, 30, 31, 52, 57] zerbitzari-bakarra kontsideratzen dutenak eta [27, 29, 97] ilara klase-anitzak kontsideratu dituztenak. Tesi honetan landu diren ereduekin lotura handiagoa duten ikerketa-artikuluek kontrol optimoa aztertzen dute bezeroak ilara utz dezaketen kasurako, adibidez [6, 7, 8, 9, 15, 24, 40, 50, 61].

Hurrengo sekzioan bezeroen abandonua kontsideratzen duten eta tesi honetan aztertzen diren ereduen aipamen labur bat aurkeztuko da.

Ilara zerbitzari-bakar klase-anitz bat bezeroen uzteekin

Demagun K bezero klase elkar lehian ari direla baliabide bakar batengatik, eta zain edo zerbitzua jasotzen ari diren bezeroak sistema utz dezaketela. Eredu hau 1.1. Irudian adierazitako problemen multzokoa da, non $M = 1$ den. Klase bakoitzari dagokionez, iritsiera prozesua Poisson prozesu bat dela onartu da eta zerbitzuak banaketa esponentziala jarraitzen duela, uzte denborak bezalaxe. Helburua orekan batez besteko kostua minimizatzea da. Kostua bezeroak ilaran mantentzeak sortzen du alde batetik, eta bestetik bezeroak ilara uzten dutenean. Ilara zerbitzari-barkar klase-anitz hau hainbat ikerketa lanetan kontsideratu da dagoeneko, adibidez [24, 81, 80]. Zein klaseri zerbitzua eman erabakitzerakoan, planifikatzaile batek kostua miopikoki minimizatzea edo uzte tasa handiena duen klasea ez zerbitzatzea aukera dezake. Sistema osoak abantailak ikus ditzake uzte tasa handia duten bezero klaseak ez zerbitzatzetik, izan ere, ilara horiek ez dira hain erraz hazten.

Orokorrean, eredu honentzat politika optimoa karakterizatzea analitikoki ezinezkoa da. Hori dela eta, tesi honen helburua kontrol optimotik gertu dauden politikak lortzea da.

Tesi honetan bi teknika kontsideratu dira eredu hau ebazteko, lehenengo teknikak erlaxazio Langrangiarrean datza, 3. Kapituluan erabili da. Bigarren teknikak gerturatze fluidoan datza, hau 5. Kapituluan erabili da. Bi tekniken teoria orokorra 1.3.1. eta 1.3.2. Sekzioan aurki daitezke, hurrenez hurren.

Inbentario kudeaketa item galkorrekin

Demagun inbentario kudeaketarako konpainia batek K item klase ezberdin ekoizten dituela eta saldu edo galtzen diren arte gordetzen dituela. Erraztasunerako $M = 1$ kontsideratuko dugu, hau da, makina bakar batek ekoitz ditzake K item klase guztiak. k klaseko item baten eskaria iristen denean konpainiara, klase horretako itemik bada gordeta konpainiak salmenta bat gauzatuko du, klase horretako itemik ez bada gordeta ordea orduan salmenta galdu egingo da eta honek kostu estra bat dakarkio konpainiari. Demagun baita itemak galkorrak direla, eta hori iruidikatzeko uzteen eredua erabiltzen dela. Helburua itemen ekoizpena kontrolatzea da, konpainian gordeta edukitzeak, salmentak galtzeak eta abandonuek dakarten kostua minimizatzeko. Itemak galkorrak izateak, makinak item gehiegi produzitzeak galerak ekar ditzake eta hobe izan daiteke makina lotarako moduan jartzea. Soluzio optimoa karakterizatzea oso konplexua gerta daiteke eta beraz, errendimendu ona erakusten duten heuristikak garatu ditugu. Problema hau 1.2. Irudiko problemen multzoan sartzen da.

Inbentarioen kudeaketarako problema hau [90, 91] ikerketa-artikuluan kontsideratu zen baina itemak galkorrak ez diren kasurako. Eredu hau 4.3.4. Sekzioan jorratuko da.

Multi-cast **igorpena uzteekin**

Eskala handiko eta bolumen handiko eduki-banaketa eta fitxategi-elkarbanatze sareetan, banda zabaleko baliabideak murritzak dira eta efikaziaz erabili behar dira. Sare hauetako trafikoa atzerapenekiko jasanbera da (*e.g.*, software eguneratzea, bideo edukia), edukiak eskatu dituzten bezeroak multzokatu daitezke, eta edukia multi-cast moduan igor daiteke saretik. Eskari bat iristen denean sistemara, gerta daiteke hobe izatea edukiaren igorpena atzeratu eta eduki berbera eskatzen duen beste bezero bat edo gehiago iritsi arte itxaron, horrela sistemak edukia bezero guztiei aldi berean igor diezaieke. Honek transmisio kapazitatea aurrezten du. Helburua, sistemak edukia igortzeko une egokia aurkitzea da, edukiaren zain dauden bezeroei zerbitzua une berean eskainiz mantentze eta zerbitzua pizte kostua minimizatzeko. Multi-cast transmisio bat martxan jartzea oso garestia gerta daiteke, helburua kostu horren eta bezeroen uzteek dakarten kostuaren oreka bat aurkitzea da. Uzte denborak banaketa esponentzial bat duela onartu dugu.

Problema hau bandit bakarreko RBP bat bezala ikus daiteke, non bandit baten egoerak eduki baten eskaria adierazten duen. Trantsizio tasak hurrengoak dira: $q^1(m, 0) > 0$, $m > 0$ bandit-a aktibatu bada eta $q^0(m, 0) = 0$, $m > 1$ bandit-a pasiboa bada, $q^a(m, m + 1)$, $m \geq 0$ trantsizioa eskari berri bati dagokio. Problema hau 1.2.2. Sekzioan azaldutako problemen kasu berezi bat da. Eredu hau 6. Kapituluan aztertuko da eta politika optimoa karakterizatuko da.

1.3 Metodologia

Aurreko sekzioetan eztabaidatu den legez, tesi honetan kontsideratu diren baliabide esleipenerako ereduak orokorrean ezin dira optimoki ebatzi, edo kasu partikularretan bakarrik lor daitezke politika optimoak. Numerikoki, MEP-ak ebazteko metodoak erabil daitezke, hala nola, *value iteration* edo *policy iteration* metodoak [22, 76]. Halere, eredu hauek numerikoki ebaztean hainbat arazo aurki daitezke: trantsizio edo egoera aldaketa tasak bornatuak ez egotea, dimentsioaren madarikazioa, egoera-espazio infinitua, etab.

Sekzio honetan MEP-ak ebazteko tesi honetan erabili diren teknikak azalduko ditugu. Helburua, 1.2.1. Sekzioan aurkeztutako RBP-entzat heuristikak lortzea da.

1.3.1. Sekzioak erlaxazio Lagrangearra aurkeztuko da, metodo honek *multi-armed* RBP-a ebazteko errazagoa den *single-armed* RBP-az gerturatzen du. 1.3.2. Sekzioan fluidoen bidezko hurbilketa matematikoa aurkeztu da, honek RBP estokastikoa, determinista den problema batez gerturatzea ahalbidetzen du. Hurbilketa metodo hauek azaldu ondoren, 1.3.3. eta 1.3.4. Sekzioetan kontrol problema estokastiko eta fluidoak optimoki ebazteko erabil daitezken tresnei buruzko eztabaida bat aurkeztu da. Tresna horiei buruzko informazio gehiago A. Eranskinean aurki daiteke.

1.3.1 Erlaxazio Lagrangearra

Sekzio honetan RBP-entzak heuristikak lortzeko metodoak aurkeztuko dira 1.2.1. Sekzioan aurkeztu den RBP-arentzak, (1.2.1) funtzio objektiboak eta erabilgarriak diren baliabideen gaineko (1.2.2) baldintzak definitzen dutena. Errendimendu ona erakusten duten heuristikak aurkitzeko ildoan, Whittle ikertzaileak [98] ikerketa-artikuluak, baliabideen gaineko (1.2.2) baldintzak erlaxatzea proposatu zuen, horrela (1.2.2) baldintza batez beste bete behar izatea eskatuz edozein denbora unitatetan bete ordez. Hau da, (1.2.2) balditza

$$\limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\left(\int_0^T\sum_{k=1}^K S_k^\phi(\vec{N}^\phi(t))\right)\leq M, \tag{1.3.1}$$

baldintzara erlaxatu daiteke.

(1.2.1) funtzio objektiboak eta (1.3.1) baliabideen gaineko baldintza erlaxatuak, *optimizazio problema erlaxatua* izenez deituko dugun problema osatzen dute. Optimizazio problema erlaxatua horrela ere adieraz daiteke: minimizatu

$$\limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\Big(\int_0^T\Big(\sum_{k=1}^K C_k(N_k^\phi(t),S_k^\phi(\vec{N}^\phi(t)))+W(\sum_{k=1}^K S_k^\phi(\vec{N}^\phi(t))-M)\Big)\mathrm{d}t\Big). \tag{1.3.2}$$

Azken hau, optimizazio problema mugatuak ebazteko oso erabilia den metodo bat da, *erlaxazio Lagrange-*iarrez ezaguna dena [20, 21], non W Lagrange biderkatzailea den. Sistemaren baliabideen gaineko baldintza funtzio objektiboan sartuz biderkatzaile batekin, lortzen den problema mugatu gabeko optimizazio problema bat da. Lortu berri den optimizazio problema RBP originala baino errazago ebatzi daiteke. (1.3.2) problema *mugatu gabeko optimizazio problema* izenez deituko dugu. W biderkatzailearen funtzioa, (1.3.1) baldintza batez beste betetzen ez duten politikei penalizazio estra bat gehitzea da.

(1.2.2) baldintzaren erlaxazioa eta Lagrange biderkatzaileen erabilera (1.3.2) funtzio objektiboa lortzeko teknika, *erlaxazio Lagrangiar* izenez ezagutzen da. Ondoren Whittle ikerlaria ohartu zen (1.3.2) problema K azpiproblemetan banatu zitekeela, bat k bandit bakoitzarentzat, hau da, K problema unidimentsionalak independenteki aztertu daitezkeela,

$$\limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\Big(\int_0^T\Big(C_k(N_k^\phi(t),S_k^\phi(N_k^\phi(t)))+WS_k^\phi(N_k^\phi(t))\Big)\mathrm{d}t\Big), \tag{1.3.3}$$

non $k\in\{1,\ldots,K\}$. RBP testuinguruan, W biderkatzailea aktibo akzioaren aurkako penalizazio gisa ikus daiteke (berdin, akzio pasiboaren laguntza gisa). Demagun $W_k(N_k)$, N_k-n (k bandit-aren egoeran) aktibo eta pasibo akzioak parekatzen dituen (1.3.3) problema optimoki askatzen duen W biderkatzailearen balioa dela, *Whittle indize* izenez ezaguna dena. Orduan, Whittle-k frogatu zuen mugatu gabeko (1.3.2) optimizazio problemaren soluzio optimoa indize motako politika zela, $W_k(N_k)$ indizetzat hartuz, indize gaitasuna delako propietatea betetzen bada. Hau da, soluzio optimoa $W_k(N_k)\geq W$ betetzen duten k bezero klase guztiak aktibo bihurtzea da optimoa, ikusi [98].

Whittle-k soluzio honetan oinarrituz heuristika bat definitu zuen (1.2.1)-(1.2.2) optimizatio problema mugatu originalerako, Whittle-en indize politika izenez ezaguna dena, eta erabakiak hartu behar diren uneoro $W_k(N_k)$ handiena duten M bandit aktibo egiten duena. Erlaxazio Lagrangearra, indize gaitasuna eta Whittle indizeari buruzko informazio gehiago 2. Kapituluan aurki daiteke.

Hainbat ikerketa-artikulutan frogatu da Whittle indize politikak oso errendimendu ona duela, ikusi [70] eztabaida baterako, eta hainbat baldintza betez gero asintotikoki optimoa dela, ikusi [59, 73, 92, 95]. Azken honek azaltzen du azken urteotan Whittle indizea kalkulatzeari ikerlariek emandako garrantzia. Adibide gisa, antolamendu oportunista hari-gabeko sareetan [1, 85, 86], webgune-*morphing*-a eta froga farmazeutikoak [47, Chapter 6]. [11]-n Whittle indize politika erabili da eduki galkorrak berreskuratzeko, [72]-n ezkutatuta dauden itemak ehizatu ahal izateko politikak lortzeko. Beste hainbat problemek ere motibatu dute hurbilketa hau, ikusi [5, 65, 66, 77, 84]. [51] ikerketa-artikulua erreferentzia ona da indize politiken aplikazioak ikusteko.

1.3.2 Fluidoen bidezko hurbilketa

Sekzio honetan prozesu estokastikoa prozesu determinista batez nola hurbil daitekeen azalduko da, jatorrizko prozesu estokastikoaren batez besteko norabidea soilik kontutan hartuz. Tesi honetan, hurbilketa hau jaiotza-eta-heriotza motako RBP-etan aplikatuko da. Demagun jaiotza tasa $b_k^a(m) := q_k^a(m, m+1)$ dela non $a \in \mathcal{A}_k$ eta $m \in E_k$, eta heriotza tasa $d_k^a(m) := q_k^a(m, m-1)$, non $a \in \mathcal{A}_k$ eta $m \in \{1, 2, \ldots\}$. Demagun $m_k(t) \in \mathbb{R}^+ \cup \{0\}$ bandit k-ri dagokion *fluido* kopurua dela. Orduan, $(N_1(t), \ldots, N_K(t))$ prozesu estokastikoa, $(m_1(t), \ldots, m_K(t))$ prozesu fluidoz hurbil daiteke non $m_k(t)$

$$\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = b_k^a(m_k(t)) - d_k^a(m_k(t)),$$

ekuazio diferentzialak deskribatzen duen edozein $k \in \{1, \ldots, K\}$-rako. Hurrengo hipotesia onartuko da tesian zehar $b_k^a(m_k)$ eta $d_k^a(m_k)$ funtzio jarraituak dira $m_k \in \mathbb{R}^+ \cup \{0\}$-n. Ez hori bakarrik, $C_k(m_k, a)$ kostu funtzioa edozein $k \in \{1, \ldots, K\}$-rako jarraitua izango da $m_k \in \mathbb{R}^+ \cup \{0\}$-n.

Hurbilketa honetatik lortzen den eredu determinista kontrol optimoaren markoan sartzen da, eta beraz, jatorrizko prozesu estokastikoa baino errazagoa da. Kontrol optimoko problemak ebazteko erabiltzen diren ohiko teknikak 1.3.4. Sekzioan azalduko dira. Kontrol problema deterministak jatorrizko optimizazio problemarekin alderatuz, propietate batzuk gal ditzake, hala ere, jatorrizko problemaren errendimenduaren edota politika optimoei buruzko ideia esanguratsuak ekar ditzake.

$(m_1(t), \ldots, m_K(t))$ prozesu determinista doitze fluido baten emaitza denaren itxaropena dugu. Hau da, denbora azeleratuz, prozesua bera denboraren azelerazioaren arabera doituz, batek prozesu limite determinista bat lortzea itxaron dezake. Tesi honetan $m_k(t)$ edozein $k \in \{1, \ldots, K\}$-rako hurbilketa huts gisa aurkeztu da eta ez da inongo ziurtasunik eskaini doiketa fluido baten emaitza den ala ez.

Hurbilketa fluidoa prozezu estokastiko bat gerturatzeko oso teknika erabilia da, eta jatorria Avram et al. [14]-en etaand Weiss [96]-en aurki daiteke. Harrigarria da nola batzuetan problema determinista ebatziz lortzen den soluzioa optimoa den jatorrizko problema estokastikoarentzat. Ikusi adibidez [14] non hau frogatu den $c\mu$ araua delakoarentzat ilara zerbitzari-bakar klase-anitz batetarako eta [18] non Kimov-en erregelarako erakutsi den *feedback* ilara klase-anitz batentzat. Beste hainbat kasuetan problema fluidoaren kontrol optimoa ez dator problema estokastikoaren optimoarekin bat, baina heuristika gisa erabil daiteke, ikusi adibidez [12, 13, 32]. Kasu horretan, asintotikoki optimoa izatea froga daiteke, hau da, kontrol fluidoa optimoa dela optimizazio problema estokastikorako doiketa apropos baten ondoren, ikusi [17, 43, 67, 68, 93]. Azkenik, aipatu behar da hurbilketa honen ospea egonkortasun propietateari dagokiola, hau da, problema fluidoa denbora finituan agortzen bada (oreka puntura iristen bada), orduan problema estokastikoa egonkorra dela esan daiteke [36, 78].

1.3.3 Kontrol estokastiko optimoa

RBP-a ebazteko kontrol estokastikoko optimoko teknikak erabiliko dira. Sekzio honetan eredu orokor baterako azalduko dira, *i.e.,* izan bedi $C(m, a)$ denbora unitateko kostua eta $q^a(m, \tilde{m})$ trantsizio tasa edozein $m, \tilde{m} \in E$ eta $a \in \mathcal{A}$-tarako, non E egoera espazioa den eta $\mathcal{A}$ akzio espazioa. Demagun $E \subseteq \mathbb{N}^d$, non $d \in \mathbb{N}$. Helburua epe-luzerako batez besteko kostuaren itxaropena minimizatzea da:

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T C(N^\phi(t), S^\phi(N^\phi(t)))\mathrm{d}t\right). \tag{1.3.4}$$

Hipotesi arin batzuk onartuz, eredu hauentzako soluzio optimoa politika egonkor bat dela froga daiteke [53]. Horregatik, politika egonkorretan oinarrituko gara. Halere, soluzio optimoa aurkitzea ez da zeregin erraza eta analisia kasuz kasu egin behar da.

Demagun erabakiak trantsizio denboratan hartzen direla. Orduan, trantsizioen arteko denborak distribuzio esponentziala jasotzen duenez, denbora jarraian dagoen jatorrizko MEP-a denbora diskretuan dagoen MEP baten baliokidea da. Azken hurbilketa hau uniformizazio teknika erabiliz egin daiteke. Horretarako, demagun trantsizio denbora m egoeran $\tau(m,a) := \sum_{\tilde{m}\in E} q^a(m,\tilde{m})$, dela, non $\tau - k$ banaketa esponentziala jarraitzen duen eta uniformeki bornatua dagoen, *i.e.,* $\tau(m,a) \leq b$ edozein $m \in E$ eta $a \in \mathcal{A}$-tarako. Soluzio optimo bat existitzeko baldintza nahikoak Bellman-en ekuaziotik lor daitezke. Hau da, $\tilde{g}$ eta $V(\cdot)$ existitzen dira non

$$\tilde{g} + \tau(m,a)V(m) = \min_{a\in\mathcal{A}}\{C(m,a) + \sum_{\tilde{m}\in E} q^a(m,\tilde{m})V(\tilde{m})\}, \tag{1.3.5}$$

$\tilde{g}$ denbora unitateko batez besteko kostu optimoa den, eta $V(\cdot)$, Balio Funtzio izenez ezagutzen denak, m egoeran eta erreferentzi egoera arbitrario baten hastearen arteko kostu ezberdintasuna jasotzen duen. $1/b$-z biderkatuz hurrengoa lor daiteke

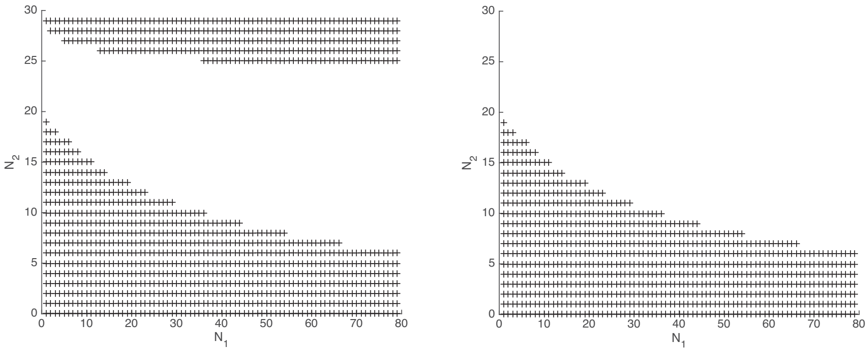
$$\frac{\tilde{g}}{b} + V(m) = \min_{a\in\mathcal{A}}\left\{\frac{C(m,a)}{b} + \sum_{\tilde{m}\in E}\frac{q^a(m,\tilde{m})}{b}(V(\tilde{m}) - V(m)) + V(m)\right\}.$$

Gogoratu $\sum_{\tilde{m}\in E} q^a(m,\tilde{m}) = \tau(m,a)$ dela, eta $g := \tilde{g}/b$, $p^a(m,\tilde{m}) = \frac{q^a(m,\tilde{m})}{b}$ trantsizio probabilitateak $\tilde{C}(m,a) = C(m,a)/b$ kostu funtzioa definitu ondoren, hurrengo ekuazioa lor daiteke

$$g + V(m) = \min_{a\in\mathcal{A}}\{\tilde{C}(m,a) + \sum_{\tilde{m}\in E} p^a(m,\tilde{m})V(\tilde{m})\}. \tag{1.3.6}$$

g balioa trantsizio bakoitzeko batez besteko kostu optimo gisa uler daiteke. (1.3.6). ekuazioa Bellman ekuazio izenez ezagutzen da eta jatorrizko MEP-aren bertsio diskretua da.

$\tau(m,a)$ trantsizio tasak uniformeki bornatuak ez daudenean, uniformizazio metodoa ezin erabil daiteke sistema diskretu egiteko. Arazo honi aurre egiteko, egoera espazioa finitua den espazio batez hurbil daiteke, espazio honi moztua deituko diogu, eta $L \in E$ parametroak mozten du, hau da, L-tik gorako tantsizioak debekatzen dira. Moztutako espazioan trantsizioak bornatuak badira, orduan uniformizazioa erabil daiteke. Hori da 1.3. irudian erabilitako metodoa. 1.3. irudian (ezkerrean) espazioa mozteko erabili den parametroa L, txikia da, ordez, 1.3. irudian (eskubian) L handi bat kontsideratu da. Ikus daitekeen bezala L parametro txikia hartzeak muga efektua izenez ezagutzen den fenomenoa sortarazten du. Efektu hau desagertuz doa L handitzen doan arabera. Analitikoki, garrantzitsua gerta daiteke Balio funtzioaren egituraren propietateak lortzea, hau da, ganbiltasuna/ahurtasuna edo modularra/azpibatukorra. Propietate hauek beharrezkoak gerta daitezke politika optimoak karakterizatzeko. Testuinguru honetan, lehenago azaldu den mozketa teknika ez da gomendagarria, izan ere, muga efektua dela eta egitura pro-



Irudia 1.3: *Value iteration* bidez lortutako emaitza zerbitzari-barkardun bi-klasetako ilara batentzat, bezeroak abandonatzen dutenean. “+"zeinua duen eremuan 1 klaseko bezeroak dute lehentasuna eta zeinurik gabeko eremuan 2 klaseko bezeroak. Ezkerreko irudian muga efektua ikus daiteke $L_1 = 80, L_2 = 30$ trunkazio parametroetarako. Eskuineko irudian ez dago muga efekturik eta mozte parametroak $L_1 = L_2 = 100$ dira.

pietateak gal daitezke. [25] ikerketa artikuluan *Smoothed Rate Truncation* (SRT) metodoa aurkeztu da. Metodo hau egoera espazio infinitua moztean datza L parametro batez, eta garrantzizkoak diren trantsizioak txikiagotzean, hau da, $q^a(m,\ell) = 0$ non $\ell \in \{m \in E : \exists i \in \{1,\dots,d\}, m_i > L_i\}$ eta $m \in E$. Baldintza batzuk betetzen direla ziurtatu ostean, froga daiteke sistema moztuaren balio funtzioak, V^L-k, $V^L \to V$ betetzen duela $L \to \infty$-rantz doanean eta egitura propietate berdinak dituela.

(1.3.6) ekuazioa Programazio Dinamikoa (PD) erabiliz ebatz daiteke, teknika honi buruzko informazioa [19, 22, 76, 79] artikuluetan aurki daiteke. Ondoren DP problemak ebazteko erabiltzen den metodoetako bat aurkeztu da, *Value iteration* izenez ezagutzen dena. Teknika hau MEP-k ebazteko erabiltzen da politika optimo egonkorrak aurkitzeko.

Izan bedi $\epsilon > 0$,

- *0 Pausua*: aukeratu $V_0(m) = 0$ edozein $m \in E$.
- *1 Pausua*: definitu
$$V_{t+1}(m) = \min_{a\in\mathcal{A}}\{\tilde{C}(m,a) + \sum_{\tilde{m}\in E} p^a(m,\tilde{m})V_t(\tilde{m})\}.$$
- *2 Pausua*: Baldin
$$\max_{m\in E}\{V_{t+1}(m) - V_t(m)\} - \min_{m\in E}\{V_{t+1}(m) - V_t(m)\} \leq \epsilon,$$
orduan $\varphi(m) = \arg\min_{a\in\mathcal{A}}\{\tilde{C}(m,a) + \sum_{\tilde{m}\in E} p^a(m,\tilde{m})V_t(\tilde{m})\}$. Horrela ez balitz, gehitu t-ri 1, $t+1$, eta *1 Pausura* joan.

Value iteration algoritmoak konbergitu dezan hainbat kondizio bete behar dira, ikusi [76]. Konbergitzen badu orduan $V_t(m) \to V(m)$ non $t \to \infty$, $V_{t+1}(m) - V_t(m) \to g$, eta $\varphi(m)$ politika optimo egonkor batera konbergitzen du.

Value iteration algoritmoa bi helburu ezberdinekin erabili da tesi honetan. Lehenengo helburua, numerikoki batez besteko politika optimoa lortu ahal izatea da. Honek, numerikoki lortutako soluzioa optimoa tesian garatu diren heuristikekin erkatzea ahalbidetzen du. Bigarrenik, *Value iteration* algoritmoa politika monotonoak optimoak diren ala ez frogatzeko erabiliko da, hau 3.2.1. Sekzioan eta 6.3. Sekzioan egingo da. Kasu batzuetan froga daiteke *value iteration* algoritmoa erabiliz politikak atari motakoak direla adibidez.

1.3.4 Kontrol optimo determinista

Sekzio honetan ekuazio diferentzialez bidez deskribatzen den sistema determinista baten kontrol dinamiko optimoa ebazteko teknikak aztertu dira. Problema hauek kontrol optimoaren teoria arloan kokatzen dira. Kontrol optimoaren teoria urte askotan erabilia izan da matematika eta ingeniaritza arloetan sortutako problemak ebazteko.

Aurrerago kontrol optimoko problema baten formulazio matematikoa aurkeztuko da eta soluzio bat optimoa izan dadin baldintza beharrezko eta nahikoak azalduko dira, ikusi A Eranskina. Enuntziatu eta frogapen osoagoak aurki daitezke [22, 34, 55, 89] liburuetan.

Kontrol optimoko problema baten formulazioa

Kontrol optimoko problema batean bi motatako aldagaiak aurki daitezke, *egoera aldagaiak* eta *kontrol aldagaiak*, azken honek lehena aldatu dezake. Helburua

$$\int_0^T C(m(t), s(t))\, \mathrm{d}t, \tag{1.3.7}$$

funtzio objektiboa minimizatzen duten kontrol ibilbidea $s(t)$ eta egoera ibilbidea $m(t)$ aurkitzea da edozein $t \in [0, T]$. Azken ekuazio honetan $C(\cdot, \cdot)$ kostu funtzioa da eta diferentziagarria da $m(\cdot)$ aldagaiarekiko, jarraitua $s(\cdot)$-rekiko eta T amaierako denbora da, non askea (optimizatu daitekeena) edo finkoa izan daitekeen. Era berean, $m(t) \in \mathbb{R}^d$, non $d \in \mathbb{N}$ eta $s(t) \in \mathcal{S}$, eta $\mathcal{S}$ kontrol onargarrien multzoa den.

Dinamika ekuazio diferentzialen bidez ematen da, hau da,

$$\frac{\mathrm{d}m(t)}{\mathrm{d}t} = f(m(t), s(t)), \text{ non } t \in [0, T], \tag{1.3.8}$$

eta non hasierako egora $m(0) = m_0$ den.

Kontrol optimoko poblemetan mota askotako egoera eta kontrol baldintzak ager daitezke, hala ere RBP-en testuinguruan kontsideratuko diren baldintzak mota honetakoak dira

$$h_1(s(t)) \leq 0, \text{ and } h_2(m(t)) \leq 0, \text{ for all } t \in [0, T], \tag{1.3.9}$$

non $h_1(\cdot)$ kontrolari soilik dagokion baldintza den, *i.e.*, kontrol aldagaia $s(t)$-ren araberakoa da, eta $h_2(\cdot)$ egoerari soilik dagokion baldintza den, egoera aldagaia $m(t)$-ren araberakoa da. Bestalde, $h_2(\cdot)$ lehenengo mailako baldintza da, hau da,

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}h_2(m(t))}{\mathrm{d}t}\right) = \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathrm{d}h_2(m(t))}{\mathrm{d}m(t)} f(m(t), s(t))\right) \neq 0.$$

$(m(t), s(t))$-ri bikote onargarria deituko diogu (1.3.9) baldintza betetzen denean eta $s(t) \in \mathcal{S}$. Ez hori bakarrik, $(m(t), s(t))$ bikote optimoa dela diogu (1.3.7). Ekuazioa globalki minimizatzen denean.

Kontrol bat optimoa izan dadin baldintza beharrezko eta nahikoak aztertzen dira hurrengo sekzioetan. Prontyagin-en Minimoaren Printzipioa (PMP), eta Hamilton-Jacobi-Bellman-en (HJB) ekuazioa azalduko dira. PMP-k optimoa izateko beharrezko baldintzak eskaintzen ditu eta HJB ekuzioak baldintza nahikoak.

Optimoa izateko beharrezko baldintzak

PMP, L.S. Pontryagin matematikari errusiarrari zor zaio. Honek ikerketa berritzailea egin zuen kontrolaren inguruan eta gaur egungo optimizazio teoriaren oinarriak ezarri zituen, ikusi [75]. PMP-k beharrezko baldintzak eskaintzen ditu $(m(t), s(t))$ bikote bat optimoa izan dadin. Egoera aldagaia eta kontrol aldagaia linealki agertzen direnean, eta egoeraren eta kontrolaren gaineko baldintzak *errazak* direnean, PMP erabil daiteke konstruktiboki *muturreko soluzio*-ak lortzeko. Muturreko soluzio izenez deitzen zaie beharrezko baldintzak betetzen dituzten $(m(t), s(t))$ bikote guztiei. Beharrezko baldintzak onargarriak diren $(m(t), s(t))$ bikoteen aukerak murrizten dituzte baina ez dute soluzio optimorik ematen.

PMP aurkeztu aurretik kontrol problema baten Hamiltondarra eta Lagrangearra definituko dira. Hamiltondarra honela definitzen da:

$$\mathcal{H}(m(t), s(t), \gamma(t)) := C(m(t), s(t)) + \gamma^{\mathsf{T}}(t) f(m(t), s(t)), \tag{1.3.10}$$

non $\gamma(\cdot)$ bektore adjuntua den. Baldintza gehigarriak badaude, (1.3.9). ekuazioan azaldutakoak bezala, Lagrangearra ere defini daiteke:

$$\mathcal{L}(m(t), s(t), \gamma(t), \nu(t), \omega(t)) := \mathcal{H}(m(t), s(t), \gamma(t)) + \nu^{\mathsf{T}}(t) h_1(s(t)) + \omega^{\mathsf{T}}(t) h_2(m(t)),$$

non $\nu(\cdot)$ eta $\omega(\cdot)$ Lagrange biderkatzaileak diren.

Kontrol optimoko problemaren Hamiltondarra eta Lagrangearra definitu ondoren, hemendik aurrera $\mathcal{H}(t)$ eta $\mathcal{L}(t)$ bezala deituko direnak, optimoa izateko beharrezko baldintzak enuntziatu daitezke $(m(t), s(t))$ bikotearentzat:

$$\begin{aligned}
&\dot{\gamma}(t) = -\frac{\partial \mathcal{L}(t)}{\partial m}, \quad \frac{\partial \mathcal{L}(t)}{\partial s} = 0, \quad s(t) = \arg \min_{s(t) \in \mathcal{S}} \mathcal{H}(t), \\
&\nu(t) h_1(s(t)) = 0, \nu(t) \geq 0 \text{ and } \omega(t) h_2(m(t)) = 0, \omega(t) \geq 0, \text{ for all } t \in [0, T].
\end{aligned}$$

A.2 Eranskinean aurki daiteke deskribapen formalago bat, tesi honen bukaeran, eta beharrezko baldintzei buruzko informazio gehiagorako [55] ikerketa artikulua gomendatzen da.

Optimoa izateko beharrezko baldintzak 5. Kapituluan erabiliko dira abandonuak gerta daitezken zerbitzari bakarreko bi-klaseko ilara baten muturreko soluzioak lortzeko.

Optimoa izateko baldintza nahikoak

HJB ekuazioa kontrol optimoaren teoriako emaitzarik garrantzitsuenetakoa da. Demagun $V(t, m)$ funtzio jarrai eta diferentziagarria existitzen dela non

$$0 = \min_{s \in \mathcal{S}} [C(m, s) + \nabla_t V(t, m) + \nabla_m V(t, m)^{\mathsf{T}} f(m, s)], \text{ for all } t, m, \tag{1.3.11}$$

non $\nabla_t V(t,m)$ eta $\nabla_m V(t,m)$ t-rekiko eta m-rekiko deribatu partzialaren adierazgarri diren, hurrenez hurren. Orduan, $V(t,m)$ balio funtzioa (1.3.7)-(1.3.9) kontrol optimo problemaren kostu optimoa da, eta (1.3.11)-ren eskuineko espresioa minimizatzen duen kontrol $s^*(m)$ ibilbidea optimoa da. Enuntziatu formalago bat A.3 Eranskinean aurki daiteke. (1.3.11) ekuazioa, HJB ekuazioa izenez ezagutzen da. Baldintza nahikoei dagokien teorema eta frogapena [22, Prop. 3.2.1] liburuan aurki daitezke.

Kasu gehienetan optimoa izateko baldintza nahikoek ez dute soluziorik aurkitzen laguntzen, izan ere, dimentsio altuko espaziotan HJB ekuazioa ezin ebatzi daiteke. Halere, optimoa izateko beharrezko baldintzen onargarriak diren soluzioen multzoa murritz dezake eta HJB erabil daiteke soluzio bat optimoa den ala ez ikusteko.

1.4 Tesiaren laburpena

Sekzio honetan tesiaren ekarpen nagusiak azalduko dira. Lehenengo kapituluak tesian zehar erabili diren ereduak azaltzen ditu eta eredu horiek nola ebazten diren erakusten du.

Tesiaren gainerako ekarpenek hurrengo egitura jarraitzen dute. I. Atalean RBP-en marko orokorra azalduko da *bandit*-ek jaiotza-eta-heriotza motako dinamika jarraitzen duten kasurako. 2. Kapituluan indize gaitasuna eta erlaxazio Lagrangearra aurkeztuko dira eta Whittle indizearen espresioa esplizitua lortuko da. Espresio hau oreka-egoerako probabilitateekiko adierazten da, non probabilitate hauek jakinak diren jaiotza-eta-heriotza prozesuentzako. Whittle indizearen errepresentazio honek indize gaitasuna oreka-egoerako probabilitateen propietate gisa definitzea ahalbidetzen du. 3. Kapituluan, 2. Kapituluko emaitzak aplikatu dira uzteak gerta daitezkeen ilara baten kasurako. Politika monotonoak optimoak direla frogatu da problema baldintza-gaberako eta baita *bandit*-ak indize gaitasuna dutela. Azken honek Whittle-en indizearen espresioa lortzea ahalbidetzen du, zeina kasu partikularretan, espliziTuki lor daitekeen. Bestalde, bezeroen uzteak indize gaitasuna ez duten problementzako indizeak berreskuratzen lagun dezake, uzteak 0-rantz eramanez. Azken hau da M/M/1 ilararen kasuan indizea lortzeko erabili den teknika kostuak ganbilak diren kasurako. Whittle indizea asintotikoki optimoa dela frogatu da trafikoa geldoa eta trafikoa arina den kasuetan, eta numerikoki ikusi da Whittle indize politikaren errendimendua ona dela trafiko intentsitate ezberdinetarako. Whittle indize politika beraz optimotik gertu dagoela ondorioztatu daiteke uzteak dituen ilara baterako. 2. eta 3. Kapituluetan aurkeztutako emaitzak [SR2] eta [SR5] artikuluetan aurki daitezke.

II Atalean baliabideen esleipenerako problema klase-anitza errazagoak diren problema deterministen bidez hurbildu dira. Bi hurbilketa matematiko kontsideratu dira. 4. Kapituluan I Atalean kontsideratutako erlaxazioa abiapuntutzat hartuz, problema zerbitzari-anitz klase-anitzaren bertsio erlaxatua doitu dugu klase bakoitzak jaiotza-eta-heriotza motako trantsizioak jasaten dituen kasurako. Doiketa honek problema determinista baldintza gabe bat lortzea ahalbidetzen du, non W Lagrange biderkatzailea indize politika bat lortzeko erabil daitekeen, indize honi *indize fluidoa* deituko diogu. Indize hau lorturik, Whittle indizeari dagokion teoria analogoa garatu daiteke eta indize fluido politika heuristika proposatu. heuristika honen Whittle-en indize politikarekiko hainbat abantaila ditut: esplizituki lor daiteke, *indexablility* propietatea erraz froga daiteke, eta kontrol fluido optimoa monotonoa dela frogatzea ere erraza da. Heuristika honen aplikagarritasuna 1.2.2. Sekzioan aurkeztutako hainbat adibideen bitartez erakutsi da. Heuristika honen errendimendua ona dela erakutsi da hainbat trafiko intentsitate ezberdinetarako eta zenbait kasuetan Whittle indizearekin bat datorrela ikusi da. 4. Kapituluan aurkeztutako emaitzak [SR3] eta [SR6] artikuluetan oinarritzen dira.

5. Kapituluan, 3. Kapituluan kontsideratutako abandonuak dituen ilara klase-anitza aztertuko da. Halere, kasu honetan jatorrizko problema estokastikoa eredu fluido batez hurbilduko da. Analisia bitan banatzen da, trafiko arinaren kasua eta trafiko geldoaren kasua. Trafikoa geldoa deneko kasuan politika optimoa guztiz karakterizatzea lortu da, politika hau lehentasun zorrotzeko politika bat da. Trafikoa arina deneko kasuan bi-klaseko ilararen kasua ebatzi da PMP erabiliz eta soluzio honetan oinarrituz heuristika bat diseinatu da klase kopuru arbitrarioa den kasurako. Bi-klaseko ilararen soluzio optimoa atari motako politika bat da. Esperimentu numerikoek erakutsi dute proposatutako heuristikak jatorrizko eredu estokastikoaren propietateak ondo jasotzen dituela. 5. Kapituluan aurkeztutako emaitzak [SR4] artikuluan oinarritzen dira.

III Atalean ere abandonuen fenomenoa aztertu da, baina kasu honetan ilaren dinamika ez da jaiotza-eta-heriotza motakoa baizik eta sorta irteera motako dinamikak aztertu dira. Problema hau aztertzeko motibazioa edukien banaketa problematik dator. Problema honetan eduki batentzako eskariak multzokatuak izan daitezke *multi-cast* moduan transmitituak izan daitezen. *Multi-cast* transmisio modua aztertuko dugu abandonuak dituen ilara klase bakar eta kapazitate infinituko zerbitzari baten kasurako. Kapitulu honetan atari politikak optimoak direla frogatu da eta ilaren luzeren banaketa kalkulatu da bi kasu ezberdinetarako: (1) zerbitzu tasa infinitua denean. eta (2) zerbitzu tasa finitua denean. Honek transmisio politika optimoa karakterizatzea ahalbidetzen du (atari motakoa dena). Numerikoki ikusi da zein den atari politika optimoa kalkulatzearen garrantzia, izan ere, optimoak ez diren atari politiken errendimendua kaxkarra da. III Atala [SR1] artikuluan oinarrituta dago.

Azkenik, A Eranskinean, soluzio bat optimoa dela frogatzea ahalbidetzen duten emaitza garrantzitsuenen bilduma aurkeztu da.

Atala I

Baliabideen esleipenerako sistema estokastikoen kontrol dinamikoa

2. Kapitulua

Indize politika jaiotza-eta-heriotza motako Restless Bandit-entzat

Contents

2.1 Sarrera **19**

2.2 Eredu matematikoa **20**

2.3 Erlaxazioa eta indize gaitasuna **22**

2.4 Whittle indize politika **27**

2.5 Whittle indize politikaren errendimendua **27**

2.6 Eranskina **28**

Kapitulu honetan, 1.2.1. Sekzioan aurkeztu den RBP-en testuingurua kontsideratu da, *bandit*-ak jaiotza-eta-heriotza motako dinamika duten kasurako. Whittle-en indize politika garatuko da eta indize gaitasuna ziurtatzen duten beharrezko baldintzak aurkeztuko dira (Whittle-en indizea garatzeko beharrezkoak direnak). Heuristika honek baliabideen esleipenerako problema konplexuen kontrol politikak definitzea ahalbidetzen du. Kapitulu honen amaieran Whittle-en indize politika muturreko egoeratan asintotikoki optimoa dela frogatzeko pausu batzuk aurkeztuko dira. Kapitulu honetan garatu den marko orokorra 3. eta 4. Kapituluetan erabiliko da errendimendu ona erakusten duten heuristikak garatzeko hainbat baliabideen esleipenerako problematan, *e.g.*, uzteak agertzen diren ilara klase-anitzetan, haririk gabeko sareetan, kontzientzia-energetikoz hornitutako zerbitzari-parketan eta inbentarioa kudeaketan item galkorrekin.

2.1 Sarrera

Orokorrean, RBP baten soluzio optimoa sarrerako parametroen eta elkarren artean lehian ari diren *bandit*-en arteko funtzio konplexu bat izan daiteke. Praktikan, politika optimoak kasu partikularretarako bakarrik dira eskuragarriak. Soluzio optimoa aurkitzea posible den kasuetako bat, 1.2.1. Sekzioan eztabaidatu den legen, MABP-ak dira, zeinentzat indize politika izenez ezagunak diren estrategiak diren soluzio optimoa.

Indize politikak optimoak diren ala ez frogatzeak ikertzaile askoren aditasuna jaso du. Problema konplexu baten soluzioa egitura oso sinple batez ebatzi daiteke, nahiz *a priori* egoera espazio guztiaren menpekoa izan litekeela uste. Adibide klasiko bezala ilara zerbitzari-bakar klase-anitza aipa daiteke, bezeroak ilaran mantentzeko kostua lineala den kasurako. Problema honentzat $c\mu$ indize araua optimoa dela frogatu da, hau da, klaseen lehentasuna $c_k\mu_k$ biderkaduraren balioak definitzen du, lehentasun ordena handienetik txikienera doalarik. Politika honetan c_k, k bezero klasearen mantzentze kostutzat jotzen da eta μ_k^{-1}, k bezero klasearen batez besteko zerbitzu tasa da, [33, 45]. Halere, kostuak ganbilak direnean, zerbitzarien edukiera egoeraren menpekoa denean eta/edo bezeroak sistema utzi dezaketenean, soluzioaren egitura sinplea desagertu egiten da [5, 26, 46]. Indize politika optimoa den beste adibide bat *Shortest-Remaining-Processing-Time (SRPT)*-ena da, non indizea bezero bakoitzak behar duen zerbitzu denbora den [82].

RBP-en kasu orokorragoan, soluzio optimoa eskuratzea oso zaila gerta daiteke. Hori dela eta, errendimendu ona erakusten duten politikak garatzea da helburua, adibidez Whittle indize politika. Orokorrean, Whittle-en indizea kasuz kasu garatu behar da. kapitulu honetan, *bandit*-ak jaiotza-eta-heriotza motako prozesuak diren kasurako Whittle indizea garatuko da espresio ezplizitu bat lortuz. Horretarako hainbat propietate betetzen direla bermatu behar da, indize gaitasuna eta politika optimoak monotonoak izatea besteak beste. Jaiotza-eta-heriotza motako prozesuak hainbat aplikazioa dituzte demografian, ilaren teorian, ingeniari sistemen errendimendu analisian, izurrien analisian eta biologian besteak beste. honek motibatu du kapitulu honen funtsa.

Kapitulu honetako gainerako emaitzak hurrengo egitura jarraituko dute. 2.2. Sekzioan eredu matematikoa aurkeztuko da, 2.3. Sekzioan erlaxazio Lagrangearra sakonago aztertuko da *bandit*-ak jaiotza-eta-heriotza motako prozesuak direnean, atari politikak aurkeztuko dira eta indize gaitasuna propietatea aztertuko da. Azken bi propitate hauek ahalbidetzen dute Whittle indizearen espresioa lortzea. 2.4. Sekzioan, Whittle indize politika definituko da eta azkenik, 2.5. Sekzioan Whittle indizea muturreko egoeretan asintotikoki optimoa izatea eztabaidatuko da.

2.2 Eredu matematikoa

K *bandit*-eko baliabide esleipenerako problema estokastikoa kontsideratu da. Demagun $N_k(t) \in \{0, 1, \ldots\}$ k *bandit*-aren egoera dela t denbora unitatean, non $k = 1, \ldots, K$. Erabakiak hartzen direneko denbora unitateetan alda daiteke *bandit* baten egoera. Erabakiak hartzen diren denbora unitate bakoitzean, planifikatzaile batek *bandit* bakoitzarekiko bi akzioen artetik bat aukeratu behar du: $a = 0$ akzioa, hau da, *bandit*-a pasibo egitea ala $a = 1$, hau da, *bandit*-a aktibo egitea, gehienez M *bandit* aktibatu daitezkeela kontuan hartuz. Kapitulu honetan zehar denbora jarraian irudikatutako jaiotza-eta-heriotza motako prozesuak jarraitzen dituzten *bandit*-ak kontsideratu dira. k *bandit*-a m_k egoeran dagoenean banaketa esponentziala jarraitzen duen denbora tarte batez aldatzen da, eta $(m_k - 1)^+$ edo $m_k + 1$ egoerara alda daiteke. k *bandit*-ari dagozkion trantsizio tasak m_k-rekiko dute menpekotasuna (eta ez gainontzeko *bandit*-en egoerekikoa). Trantsizioak, *bandit*-aren egoera N_k denean edozein $k = 1, \ldots, K$ delarik, honakoak dira

$$\begin{cases} \vec{N} \to \vec{N} + \vec{e}_k & b_k^a(N_k) \text{ transizio tasarekin,} \\ \vec{N} \to \vec{N} - \vec{e}_k & d_k^a(N_k) \text{ transizio tasarekin,} \end{cases} \tag{2.2.1}$$

non $\vec{N} = (N_1, \ldots, N_K)$ den eta $\vec{e}_k$ K-dimentsioko zeroz osatutako bektore bat den, zeinak k. osagarria 1-a duen. Bestalde, $d_k^a(0) = 0$.

Ohartu *bandit* baten trantsizioak aukeratu den akzioarekiko duela menpekotasuna. Bereziki, *bandit* baten egoera akzioa pasiboa zein aktiboa denean aldatu daiteke.

ϕ politikak erabakitzen du noiz egiten den *bandit* bat aktibo. Markov propietateari esker, une bakoitzean *bandit*-aren egoerarekiko soilik menpekotasuna duten politiketan jarriko da arreta. ϕ politika bat emanda, $N_k^\phi(t)$ k *bandit*-aren egoeraren adierazle izango da t denbora unitatean eta $\vec{N}^\phi(t) = (N_1^\phi(t), \ldots, N_K^\phi(t))$. Demagun $S_k^\phi(\vec{N}^\phi(t)) \in \{0,1\}$ aldagaiak k *bandit*-a t denbora unitatean ϕ politikak aktiboa ala pasiboa egiten duen adierazten duela. K *bandit*-etatik gehienez M egin daitezke aktibo, edo baliokideki, gutxienez $K - M$ *bandit*-ek izan behar dute pasibo. Hori dela eta, hurrengo baldintza daukagu

$$\sum_{k=1}^{K} (1 - S_k^\phi(\vec{N})) \geq K - M, \tag{2.2.2}$$

zeina (1.2.2)-ren baliokidea den. Izan bedi $C_k(m,a)$ denbora unitateko k *bandit*-aren kostu funtzioa m egoeran, pasiboa ($a = 0$ akzioa) zein aktiboa ($a = 1$ akzioa) izan daiteke. Baldintza hau betetzen duten kontrol estrategiei onargarriak deituko diegu eta politika onargarrien multzoari $\mathcal{U}$.

Helburua *scheduling* politika optimoa aurkitzea da, OPT gisa adieraziko dena, epe luzera batez besteko kostuaren itxaropena minimizatzen duena

$$\mathcal{C}^\phi := \limsup_{T \to \infty} \sum_{k=1}^{K} \frac{1}{T} \mathbb{E} \left(\int_0^T C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t)))\, \mathrm{d}t \right). \tag{2.2.3}$$

Izan bedi $\mathcal{C}^{OPT} := \min_{\phi \in \mathcal{U}} \mathcal{C}^\phi$ OPT politika optimoak eragindako batez besteko kostua. Jakina da (2.2.3) problemarentzat g eta $V(\cdot)$ existitzen direla zeinak Bellmanen ekuazioa betetzen duten

$$\begin{aligned} g = \min_{\vec{s}, s.t. \sum_k s_k \leq M} \Bigg(\sum_{k=1}^{K} \Bigg[& C_k(m_k, s_k) + b_k^{s_k}(m_k) V(\vec{m} + e_k) + d_k^{s_k}(m_k) V(\vec{m} - e_k) \\ & - (d_k^{s_k}(m_k) + b_k^{s_k}(m_k)) V(\vec{m}) \Bigg] \Bigg). \end{aligned} \tag{2.2.4}$$

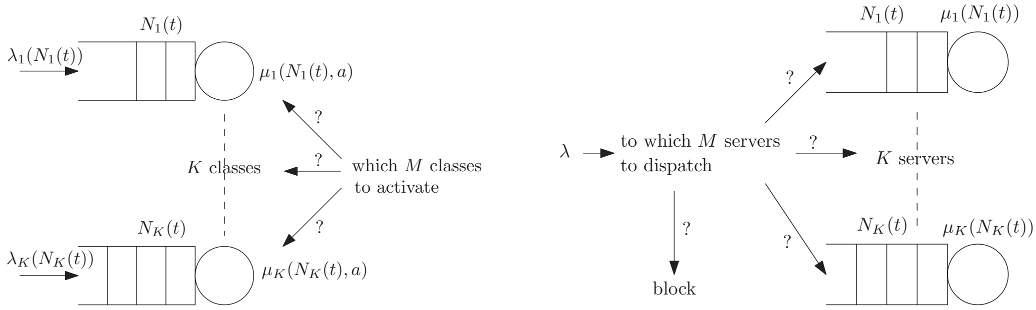
(2.2.4) problema minimizatzen duen politika egonkorra optimoa da, ikusi 1.3.3. Sekzioa. (2.2.4). ekuazioa, (1.3.5). ekuazioaren baliokidea da, K *bandit*-en trantsizioak jaiotza-eta-heriotza motakoak direnean. Orduan, $g = \min_\phi \mathcal{C}^\phi$ eta $V(\vec{m})$ balio funtzioa da.

Hurrengo sekzioan kapitulu honetan azalduko den hurbilketa matematikoaren motibazioa aurkeztuko da RBP-en markoan sartzen diren bi adibide mota azalduz.

2.2.1 Adibideak

Problema hau aztertzeko motibazioa, laburki 1.2.2. Sekzioan azaldu den legez, ilara zerbitzari-anitz klase-anitz-etan baliabideen esleipenerako problemetatik dator. Demagun K bezero klase daudela, eta *bandit* bat existitzen dela klase bakoitzaren adierazpide. N_k aldagaiak ilaran zain dauden k klaseko bezero kopurua adierazten du. Ez hori bakarrik, $b_k^a(N_k)$ eta $d_k^a(N_k)$ funtzioek iritsiera eta irteera tasak adierazten dituzte, hurrenez hurren. Egoerarekiko dependentzia duen irteera tasa izatean fenomeno garrantzitsuak irudikatzea ahalbidetzen du, hala nola, kontzientia energetikoz hornitutako zerbitzari-baserriak, bezeroen pazientziarik eza, non bezeroak sistema uzten duten zerbitzua jasotzen bukatu aurretik. Lehenengo honetan, irteera tasa $(N_k)^\alpha$ abiadura-doiketa gaiarekiko proportzionala da, ikusi [99], eta bigarren adibidean irteera tasak $\theta_k N_k$ gaia hartzen dute kontuan, non θ_k uzteen tasa den k klaseko bezeroentzak, ikusi [8, 48]. Marko honen erabilgarritasuna erakusteko, RBP-en testuinguruan sartzen diren bi problema mota orokor aurkeztuko dira. Bi problema mota hauek sakonago garatuko dira 3. eta 4. Kapituluetan (4.3. Sekzioan).

Lehenengo problema mota 2.1. Irudian (Ezkerrean) irudikatu den problemei dagokie. Helburua lehian ari diren zein M klase aktibatuko diren aukeratzea da. Beraz, transizio tasak hurrengoak dira: $b_k^a(N_k) = \lambda_k(N_k)$ eta $d_k^a(N_k) = \mu_k(N_k, a)$, non $a = 1$ den k klaseko bezeroak zerbitzatuak badira. Klase bakoitzaren iritsiera tasa klase horren ilararen luzeeraren menpekoa izatea onartuko dugu, eta irteera tasa klasearen ilararen luzeeraren zein akzioaren $a \in \{0, 1\}$ menpekoa izatea. 3. Kapituluan eta 4.3.1. Sekzioan eredu hau erabiliko da *scheduling* politikak aztertzeko abandonuak gerta daitezken ilara zerbitzari-bakar klase-anitzetarako, eta 4.3.2. Sekzioan haririk gabeko sareen *scheduling* optimoa aztertzeko erabiliko da, *scheduling* oportunista medio, zerbitzariaren edukiera handituz doa bezeroen kopurua handitzen den heinean, ikusi [28]. Bestalde, $M = 1$ eta $\mu_k(N_k, a) = \mu_k$, eredu honek ilara zerbitzari-bakar klase-anitza atzematen du.



Irudia 2.1: Left: A multi-class system where M classes can be simultaneously served. Right: Load balancing in a multi-server system

Bigarren motako problemak trafikoaren karga banaketarako problemak dira, ikusi 2.1. Irudia (Eskuinean), non iritsiera berriak K zerbitzari ezberdinetara bidaltzen diren ala blokeatu egiten diren. Iritsiera bakoitza gehienez M zerbitzarietara bidal daiteke, non $M = 1$ den balio tipikoa trafikoaren karga banaketa problemetan. Tranzisioak horrela idatz daitezke orduan: $b_k^a(N_k) = \lambda a$ eta $d_k^a(N_k) = \mu_k(N_k)$, non $a = 1$ iritsi berri den bezeroa k zerbitzarira bidaltzen bada. 4.3.3. eta 4.3.4. Sekzioetan bezeroak optimoki nola zerbitzarietara nola bidali aztertzen da, (i) kontzientzia energetikoz hornitutako zerbitzari-parke batean, non zerbitzarien edukierak abiadura-doiketa arau bat jarraitzen duen, eta (ii) inbentarioen kudeaketa itemak galkorrak direnean.

2.3 Erlaxazioa eta indize gaitasuna

Sekzio honetan erlaxazio lagrangearra deskribatuko da, [98] ikerketa artikuluan Whittle-ek proposatu zuena eta 1.3.1. Sekzioan laburki eztabaidatu dena, (2.2.3) problemaren soluzioaren hurbilketak garatzeko (2.2.3) baldintzapean. Hurbilketa matematiko emankor bat problema erlaxatua analizatzea da non aktiboak izan daitezkeen *bandit*-en gaineko baldintzak uneoro bete beharrean batez beste betetzea eskatzen den, hau da,

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \sum_{k=1}^K (1 - S_k^\phi(\vec{N}^\phi(t)))\mathrm{d}t\right) \geq K - M. \tag{2.3.1}$$

Problema erlaxatuaren helburua (2.2.3) problema ebazten duen politika aurkitzea da (2.3.1) baldintza betetzen dela bermatuz. Problema erlaxatuarentzako politika optimoa, indize motako politika dena, heuristika gisa erabil daiteke jatorrizko optimizazio problemarentzako. $\mathcal{U}^{REL}$ multzoa, (2.3.1) baldintza betetzen duten politikak adieraziko erabiliko da, eta ohartu $\mathcal{U} \subset \mathcal{U}^{REL}$.

Problema erlaxatua hurrengo problema mugatu gabea kontsideratuz ebatz daiteke: aurkitu ϕ politika

$$\begin{aligned}&\mathcal{C}^\phi(W)\\ &:= \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Bigg(\int_0^T \left(\sum_{k=1}^K C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t))) + W(K - M - \sum_{k=1}^K (1 - S_k^\phi(\vec{N}^\phi(t))))\right)\mathrm{d}t\Bigg),\end{aligned} \tag{2.3.2}$$

minimizatzen duena, non W Lagrange biderkatzailea den. Demagun W finkoa dela eta $REL(W)$ (2.3.2) problema minimizatzen duen politika dela, eta demagun $\mathcal{C}^{REL(W)}(W) := \min_{\phi\in\mathcal{U}^{REL}} \mathcal{C}^\phi$ problema erlaxatuaren errendimendu optimoaren adierazpide dela. Orduan, edozein $W \geq 0$, biderkatzailerentzako, $\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^{OPT}$ betetzen da. Hori ikusi ahal izateko, ohartu $W \geq 0$ finko baterako eta $\phi \in \mathcal{U}$

$$\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^\phi(W) \leq \mathcal{C}^\phi,$$

betetzen dela. Lehenengo inekuazioa $REL(W)$-ren definizioaren ondorioa da, eta bigarren inekuazioa $1 - K + \sum_{k=1}^K (1 - S_k^\phi(\vec{N}^\phi(t))) \geq 0$-tik dator edozein $\phi \in \mathcal{U}$.

(2.3.2) problema K azpiproblematan deskonposatu daiteke, bat k *bandit* bakoitzeko. Oreka eta ergodikotasun hipotesiak onartuz (2.3.2) problema sinplifika daiteke, hau da, aurkitu ϕ politika

$$\mathcal{C}_k^\phi(W) := \mathbb{E}(C_k(N_k^\phi, S_k^\phi(N_k^\phi))) - W\mathbb{E}(\mathbf{1}_{S_k^\phi(N_k^\phi)=0}), \tag{2.3.3}$$

minimizatzen duena, non N_k^ϕ oreka egoeran dagoen ϕ politika jarraitzen duen k *bandit*-aren distribuzio berdina duen. (2.3.2)-ren soluzioa (2.3.3) K optimizazio azpiproblemen emaitzak batuz lor daiteke.

(2.3.3) problema MEP bat bezala ikus daiteke eta soluzio optimoa Bellman ekuazioaren soluzioa da

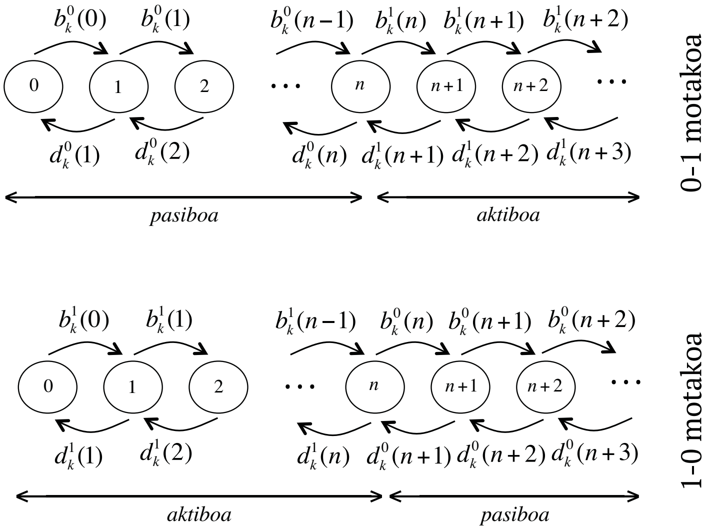
$$\begin{aligned} g_k(W) = \min\Big(&C_k(m,1) + b_k^1(m)\Delta V(m) - d_k^1(m)\Delta V(m-1),\\ &C_k(m,0) - W + b_k^0(m)\Delta V(m) - d_k^0(m)\Delta V(m-1)\Big), \end{aligned} \tag{2.3.4}$$

non $g_k(W) = \min_\phi \mathcal{C}_k^\phi(W)$ politika optimoak emandako kostu minimoa den eta $\Delta V(m) = V(m+1) - V(m)$.

Kasu partikular batzuetan (2.3.3) problemaren soluzioa monotonoa dela froga daiteke. Hurrengo sekzioan monotonoa izatearen definizioa aurkeztuko dugu eta propietate hau erraz froga daitekeen kasuak aztertu.

2.3.1 Atari-politikak

Problema batzuetan, (2.3.3)-ren soluzio optimoa atari politika motakoa dela froga daiteke. Hau da, politika monotonoak optimoak direla ikusi daiteke: atari bat existitzen da, $n_k(W)$ izendatuko duguna, zeinarentzat



Irudia 2.2: Transizioak atari-politikapean; goian 0-1 motakoak, behean 1-0 motakoak.

k *bandit* bat $m_k \leq n_k(W)$ egoeran dagoenean, akzioa optimoa a den, eta bestela a', non $a, a' \in \{0, 1\}$ eta $a \neq a'$. Izan bedi $\phi = n$ atari politika bat n atariarekin, eta 0-1 motakoa dela esango da $a = 0$ eta $a' = 1$ badira eta 1-0 motakoa dela $a = 1$ eta $a' = 0$ badira (ikusi 2.2 Irudia). RBP-en testuinguruan atari politika bat optimoa dela frogatu da adibidez [5, 48, 90] artikuluetan. Adibide gehiago aurki daitezke [47, Section 6.5] artikuluan. Hurrengo proposizioan (2.3.3)-ren soluzioa atari motako politika izateko baldintza nahikoak aurkeztuko dira. Frogapena 2.6.1 Eranskinean aurki daiteke.

2.1 Proposizioa. *Demagun* $b_k^a(N_k) = \lambda_k(N_k)$ *eta* $d_k^a(N_k) = \mu_k(N_k)a$. *Orduan, existitzen da* $n_k = -1, 0, 1, \ldots$ *non 0-1 motako atari politika,* n_k *atariarekin,* (2.3.3) *problemarentzat optimoa den. Ordez,* $b_k^a(N_k) = \lambda_k(N_k)a$ *eta* $d_k^a(N_k) = \mu_k(N_k)$ *badira, orduan existitzen da n* $n_k = -1, 0, 1, \ldots$ *non 1-0 motako atari politika optimoa den* (2.3.3) *problemarentzat.*

2.1. Proposizioko hipotesiez gain, ez da existitzen baldintza nahikorik jaiotza-eta-heriotza motako *bandit* batean politika monotonoak optimoak izatea ziurtatzen duenik. 3. Kapituluan atari politikak frogatu dira uzteak gerta daitezken ilara klase-anitz batentzat.

Hurrengo sekzioan indize gaitasuna eta Whittle indizea aurkeztuko dira. Ez hori bakarrik, atari politikak optimoak direla froga daitezken kasuan Whittle indizea esplizituki garatuko da.

2.3.2 Indize-gaitasuna eta Whittle indizea

Indize gaitasuna jatorrizko problemarentzat heuristika bat garatzea ahalbidetzen duen propietatea da. Propietate honek Lagrange biderkatzailea, edo baliokideki W pasibo akzioa bultzatzeko subsidioa handitzen den heinean pasibo akzioa agindu den egoeren multzoa haztea eskatzen du. Whittle ikertzailea izan zen lehena propietate hau definitzen [98] artikuluan. Tesi honetan hurrengo definizio formala onartuko da.

2.1 Definizioa. *Bandit batek indize gaitasuna propietatea betetzen duela esango da, pasibo akzioa* (2.3.3) *problemarentzat optimoa den egoeren multzoa (*$D_k(W)$ *izenez deituko dena) handituz badoa* W *hazten den heinean, hau da,* $W' < W \Rightarrow D_k(W') \subseteq D_k(W)$.

Indize gaitasuna izango balu, Whittle indizea N_k egoeran horrela definituko litzateke:

2.2 Definizioa. *Bandit batek indize gaitasuna badu, Whittle indizea N_k egoeran, N_k-n hartzen den akzioak* (2.3.3) *problemaren emaitza optimoarengan eraginik izan ez dezan W-k har dezakeen baliorik baxuena bezala definitzen da. Whittle indizea $W_k(N_k)$ bezala definitzen da.*

(2.3.2) problema erlaxatuaren soluzioa W pasibo izateagatik ematen den subsidioa baino Whittle indize handiagoa duten eta n_k egoeran dauden *bandit* guztiak aktibo bihurtzea izango da, *i.e.*, $W_k(n_k) > W$. Bereziki, argumentu Lagrangear batek erakusten du existizen dela $W = W^*$ balio bat, zeinentzat (2.3.1) baldintza zorrotz betetzen den, *i.e.*, (2.3.2) problemaren, ϕ soluzio optimoak $W = W^*$ balioarentzat, batezbeste M *bandit* aktibatuko ditu.

Hau esanik, Whittle indizea garatu daiteke (2.3.3) problemaren soluzioa n atari motako politika bat den kasurako. Batez besteko kostua oreka-egoeran dauden probabilitateen bidez adieraz daiteke, eta jaiotza-eta-heriotza motako prozesuentzak probabilitate hauek ezagunak dira. Demagun $\pi_k^n(m)$, m egoeraren eta k *bandit*-aren oreka-egoerako probabilitatea dela n atari politikapean. Definitu

$$g_k^{(n)}(W) := \sum_{m=0}^{\infty} C_k(m,a)\pi_k^n(m) - \begin{cases} W\sum_{m=0}^{n}\pi_k^n(m) & \text{if 0-1 motako atari politika,} \\ W\sum_{m=n+1}^{\infty}\pi_k^n(m) & \text{if 1-0 motako atari politika.} \end{cases}$$

Orduan (2.3.3) problemaren batez besteko kostu optimoa

$$g_k(W) = \min_n\{g_k^{(n)}(W)\}, \tag{2.3.5}$$

balioak ematen du.

Orain Whittle indizea lortzeko pausuak enuntziatuko dira. 2.2. Proposizioaren frogapena 2.6.2. Eranskinean aurki daiteke.

2.2 Proposizioa. *Demagun* (2.3.3) *problemaren soluzio optimo bat atari motakoa dela, eta* $\sum_{i=0}^{n}\pi_k^n(i)$ *hertsiki gorakorra dela n-n, non $\pi_k^n(m)$, m egoeraren oreka-egoerako probabilitatea den n atari politikapean. Orduan* (2.3.3) *problema indexable-a da eta k bandit-ari dagokien Whittle indize balioa hurrengo pausuak jarraituz kalkula daiteke:*

- **0. Pausua** *Kalkulatu*

$$W_0 = \inf_{n\in\mathbb{N}\cup\{0\}} \frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{-1}, S_k^{-1}(N_k^{-1})))}{\sum_{m=0}^{n}\pi_k^n(m)},$$

eta n_0 izenez deitu minimizatzaile handienari, atari-politika 0-1 motakoa bada. Kalkulatu

$$W_0 = \sup_{n\in\mathbb{N}\cup\{0\}} -\frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{-1}, S_k^{-1}(N_k^{-1})))}{\sum_{m=0}^{n}\pi_k^n(m)},$$

eta ditu n_0 izenez maximizatzaile handienari, atari-politika 1-0 motakoa bada. Orduan, definitu $W_k(n) := W_0$ edozein $n \leq n_0$. $n_0 = \infty$ bada $W_k(n) := W_0$ definituko da edozein $n > n_0$, bestela 1. Pausura egin salto.

- ***j.* Pausua** *Kalkulatu*

$$W_j = \inf_{n \in \mathbb{N}\setminus\{0,\ldots,n_{j-1}\}} \frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{n_{j-1}}, S_k^{n_{j-1}}(N_k^{n_{j-1}})))}{\sum_{m=0}^{n} \pi_k^n(m) - \sum_{m=0}^{n_{j-1}} \pi_k^{n_{j-1}}(m)}, j \geq 1,$$

eta n_j izenez deitu minimizatzaile handienari 0-1 motakoa bada atari-politika. Kalkulatu

$$W_j = \sup_{n \in \mathbb{N}\setminus\{0,\ldots,n_{j-1}\}} -\frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{n_{j-1}}, S_k^{n_{j-1}}(N_k^{n_{j-1}})))}{\sum_{m=0}^{n} \pi_k^n(m) - \sum_{m=0}^{n_{j-1}} \pi_k^{n_{j-1}}(m)}, j \geq 1,$$

eta n_j izenez deitu maximizatzaile handiena 1-0 motakoa bada atari-politika. Ostean, $W_k(n) := W_j$ definitu edozein $n_{j-1} < n \leq n_j$. $n_j = \infty$ bada orduan $W_k(n) = W_j$ edozein $n > n_j$, bestela egin salto $j+1$. pausura.

2.2. Proposizioak Whittle indizea garatzeko errezeta orokor bat eskaintzen du *bandit*-en dinamika jaiotza-eta-heriotza motako prozesu bat denean:

(i) Politika monotonoak optimoak direla frogatu.
(ii) Indize gaitasuna zehaztu, hau da, frogatu $\sum_{m=0}^{n} \pi_k^n(m)$ hertsiki gorakorra dela.
(iii) (i) eta (ii) frogatu badaitezke, orduan Whittle indizea 2.2. Proposizioak ematen du.

(i) eta (ii) pausuak ereduaren menpekoak dira. (iii) pausua berehalakoa da eta indizea beti 2.2. Proposizioak emango du. Azken prozedura hau 3. Kapituluan erabiliko da uzteak gerta daitezken ilara klase-anitz batean, eta 4. Kapituluan hainbat trafikoaren karga banaketa eta klase aukeraketa problemetan.

Hurrengo korolarioan Whittle indizea karakterizatu da $n_i = i$ den kasu partikularrerako, non $i \in \mathbb{N} \cup \{0\}$, eta n_i 2.2. Proposizioan definitu den bezala. Frogapena 2.6.3. Eranskinean aurki daiteke.

2.1 Korolarioa. *Demagun (2.3.3) problemaren soluzio optimoa atari-politika bat dela, eta $\sum_{m=0}^{n} \pi_k^n(m)$ hertsiki gorakorra dela n-n, non $\pi_k^n(m)$ k bandit-aren m egoeraren oreka-egoeraren probabilitatea den n atari-politikapean. Orduan, k bandit-aren indizea definitu daiteke.*

(2.3.3) problemaren soluzio optimoaren egitura 0-1 motako atari-politika bada, orduan,

$$\frac{\mathbb{E}(C_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(C_k(N_k^{n-1}, S_k^{n-1}(N_k^{n-1})))}{\sum_{m=0}^{n} \pi_k^n(m) - \sum_{m=0}^{n-1} \pi_k^{n-1}(m)}, \tag{2.3.6}$$

ez-beherakorra bada n-n, Whittle indizea $W_k(n_k)$ (2.3.6)-k definitzen du eta beraz, ez-beherakorra da. Modu berean, (2.3.3) problemaren soluzio optimoaren egitura optimoa 1-0 motakoa bada, orduan, (2.3.6) ez-beherakorra da n-n, $-W_k(n_k)$ (2.3.6)-k definitzen du eta beraz, Whittle indizea ez-gorakorra da.

(2.3.6) itxurari buruzko ohar bat azalduko da orain. (2.3.6)-ren zenbakitzailea n egoeran pasibo izatea aukeratzeak dakarren kostuaren handitzea bezala interpreta daiteke. Bestalde, izendatzailea pasibo izatearen tasa handitzea bezala uler daiteke, non zenbat eta pasiboago izan subsidioa handiagoa den. Hori dela eta, $W(n)$ kostuaren handitzea pasibo izatearen tasa unitatearekiko bezala uler daiteke, gai hau *Marginal Productivity Index* izenez definitu zuen Nino-Morak [70].

Tesi honen autoreek dakitenez, jaiotza-eta-heriotzako motako prozesu gisa adieraz daitezkeen *bandit*-entzak Whittle indizea esplizituki lor daitekeenik ez da publikatua izan. Honen arrazoi bat (2.3.4) problema ebazteak duen zailtasuna izan daiteke, izan ere, g_k eta $V_k(m)$ bi ezezagun ditu. Ikerlariak zailtasun hori saihestu nahian, problema deskontatua ebatzi izan ohi dute, kostu deskontatuaren Bellman ekuazioak ezezagun bakarra du, eta 2.2. Proposizioan batez besteko kostuarentzat egin den legez, deskontatutako kostuarentzat egin daiteke eta ondoren limitea hartu batez besteko emaitza berreskuratzeko. Hurbilketa hau da adibidez [5] artikuluan erabili den metodoa eta [47, Section 6.5] artikuloan bi-norabidezko *bandit*-entzat, non akzio pasibo eta aktiboek prozesua kontrako norabideetan bultzatzen duten. [48] artikuluan autoreek algoritmo bat garatu dute ilara klase-anitz batean onarpen kontrolaren indizea kalkulatzeko. Eredu guzti hauek komunean dute, erlaxatu ondoren, jaiotza-eta-heriotza motako prozesuez deskriba daitezkeela eta Whittle indizea (2.3.6)-k definitzen duela. Bi norabidezko *bandit*-ei dagokienez zuzenean froga daiteke (2.3.6) indizea eta [47, Theorem 6.4]-n lortutako indizea baliokideak direla.

2.4 Whittle indize politika

Sekzio honetan problema erlaxatuaren soluzioa nola erabil daitekeen jatorrizko problemaren soluzio bat lortzeko azalduko da. Problema erlaxatuaren soluzioa, hau da, $W_k(n_k) > W$ betetzen duten eta n_k egoeran dauden *bandit*-ak aktibo bihurtzea, soluzioa onartezina izan daiteke jatorrizko ereduarentzat non gehienez M *bandit* egin daitezken aktibo. Beraz, Whittle-ek, [98]-n, heuristika bat proposatu zuen, Whittle indize politika izenez ezagutzen dena.

2.3 Definizioa (Whittle indize politika)**.** *Demagun t denbora unitatean $\vec{N}(t) = \vec{n}$ egoeran dagoela sistema. Whittle indize politikak $W_k(n_k)$ indizearen baliorik handiena dueten M bandit-ak aktibo bihurtzea aukeratzen du.*

Ohartu Whittle-en indizea *bandit* guztientzat negatiboa bada, *bandit* guztiak pasibo mantentzen direla. Azken hau, optimizazio problema erlaxatuaren ondorio bat da: Whittle indizea negatiboa denean $\tilde{n}$ egoera batean, honek esan nahi du aktibo bilakatuko dela baldin eta soilik baldin $W < W_k(\tilde{n}) < 0$, hau da, *kostu* bat ordaintzen denean pasiboa izateagatik.

2.5 Whittle indize politikaren errendimendua

Aurreko sekzioan Whittle-en indize politika definitu da, eta orain heuristika hau optimoa den ala ez eztabaidatu daiteke. Eztabaida hau 3.5.2. Sekzioan erabiliko da Whittle indize politika optimoa dela frogatzeko muturreko eremuetan, hau da, trafikoa arina eta trafikoa geldoa den kasuetan.

Gogoratu $\mathcal{U}$ eta $\mathcal{U}^{REL}$ jatorrizko problema eta problema erlaxatuarekiko onargarriak diren politiken adierazpide direla, hurrenez hurren, eta $\mathcal{U} \subseteq \mathcal{U}^{REL}$. 2.2. Sekzioan argudiatu den legez, $W \geq 0$ biderkatzailearen edozein baliotarako $\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^{OPT}$, non $\mathcal{C}^{REL(W)}(W)$ and $\mathcal{C}^{OPT}$ problema erlaxatuaren eta jatorrizko problemaren kostu minimoak diren, hurrenez hurren. Gogoratu baita ere $\mathcal{C}^{REL(W)}(W)$ balioa lortzeko Whittle indizea W baino handiagoa duten klaseak zerbitzatuz lortzen dela. Deitu $\mathcal{C}^{WI}$ Whittle indize politika onargarriaren errendimenduari eta finkatu $\mathcal{C}^* = \sup_W \mathcal{C}^{REL(W)}(W)$. Orduan tribiala da

$$\mathcal{C}^{REL(W)}(W) \leq \mathcal{C}^* \leq \mathcal{C}^{OPT} \leq \mathcal{C}^{WI}, \tag{2.5.1}$$

betetzen dela.

Orain ohartu

(i) $REL(0) \in \mathcal{U}$, edo,

(ii) $REL(W) \in \mathcal{U}$ eta (2.3.1) baldintza berdintasunez betetzen dela,

betetzen bada, orduan W-ren aukera horretarako $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^* = \mathcal{C}^{OPT} = \mathcal{C}^{WI}$, eta kasu horietan Whittle-en indize politika optimoa da. Hori frogatzeko ohartu $REL(W) \in \mathcal{U}$, orduan $REL(W)$ politika Whittle-en indize politikaren baliokidea da. Orduan, $W = 0$ hartuz, $\mathcal{C}^{REL(0)}(0) = \mathcal{C}^{REL(0)} = \mathcal{C}^{WI}$ lortzen da, non lehenengo berdintasuna definizioz beretzen den $W = 0$ izanagatik. Demagun $W > 0$, orduan (2.3.1) baldintza berdintasunez betetzen denez, $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^{REL(W)} = \mathcal{C}^{WI}$ lor daiteke berriz. Bi kasuetan, (2.5.1) erabili da, $\mathcal{C}^{REL(W)}(W) = \mathcal{C}^* = \mathcal{C}^{OPT} = \mathcal{C}^{WI}$ betetzen dela ondorioztatzeko. [47, 6. Kapitulua] eta [48, 5. Sekzioan] artikuluetan hurbilketa berdina deskribatzen da.

(i) edo (ii) frogatu daitekeenean mugako eremuetan, Whittle indize politika asintotikoki optimoa dela frogatzen da.

2.6 Eranskina

2.6.1 2.1. Proposizioaren frogapena

Klaseekiko dependentzia, k-re bitartez, alde batera utzi da frogapen honetan.

Demagun $b^a(N) = \lambda(N)$ eta $d^a(N) = \mu(N)a$ direla, $b^a(N) = \lambda(N)a$ eta $d^a(N) = \mu(N)$ kasua antzera egin daitezke. Sistema ergodikoa dela kontsideratu denez politika egonkorretan oinarritzen da analisia, eta demagun existitzen dela ϕ^* politika egonkorra (2.3.3) problema optimoki ebazten duen. Definitu $n^* = \max\{m \in \{0, 1, \dots\} : S^{\phi^*}(m) = 0\}$, orduan trantsizio tasen definizioa dela eta $S^{\phi^*}(n^*) = 0$ eta $S^{\phi^*}(m) = 1$ edozein $m > n^*$. (2.3.3) ekuazioak emandako batez besteko kostua ϕ^* politika optimopean hurrengoa da:

$$\mathbb{E}(C(N^{\phi^*}, S^{\phi^*}(N^{\phi^*}))) - W\mathbb{E}(\mathbf{1}_{S^{\phi^*}(N^{\phi^*})=0}) = C(n^*, 0)\pi^{\phi^*}(n^*) + \sum_{m=n^*+1}^{\infty} C(m, 1)\pi^{\phi^*}(m) - W\pi^{\phi^*}(n^*)$$
$$= \mathbb{E}(C(N^{n^*}, S^{n^*}(N^{n^*}))) - W\pi^{n^*}(n^*),$$

hau da, 0-1 motako atari politikak n^* atariarekin, errendimendu optimoa du.

2.6.2 2.2. Proposizioaren frogapena

0-1 motako atari politiken kasua aztertuko da frogapen honetan, n atari politika 1-0 motakoa den kasuan antzeko moduan froga daiteke.

Demagun pausuak $J \in \mathbb{N} \cup \{\infty\}$ iterazioan gelditu dela, eta beraz, $n_J = \infty$. Bestalde, $W_i := W_J$ eta $n_i = \infty$ edozein $i \in \{J + 1, \dots\} \cup \{\infty\}$. Lehenik eta behin, $W_0 < W_1 < W_2 < \dots < W_J$, frogatuko da eta ohartu definizioz n_i non $i \in \mathbb{N} \cup \{0, \infty\}$ segida gorakorra den. Lehenik, $W_i < W_{i+1}$ frogatuko da edozein

$i \in \{0, 1, 2, \ldots, J\}$. W_i-ren karakterizazioa dela eta

$$
\begin{aligned}
&\frac{\mathbb{E}(C(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)} \\
&> \frac{\mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)} \\
&\Longrightarrow (\mathbb{E}(C(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))) \left(\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m) \right) \\
&> (\mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))) \left(\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m) \right),
\end{aligned}
$$

non azken inekuazioa $\sum_{m=0}^{n} \pi^n(m)$ hertsiki gorakorra izateak ematen duen.

$$
\mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i})))(\sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m) - \sum_{m=0}^{n_i} \pi^{n_i}(m)),
$$

gaia inekuazioaren bi aldeetan gehituz, kalkulu batzuen ostean hurrengoa lor daiteke

$$
\begin{aligned}
W_{i+1} &= \frac{\mathbb{E}(C(N^{n_{i+1}}, S^{n_{i+1}}(N^{n_{i+1}}))) - \mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i})))}{\sum_{m=0}^{n_{i+1}} \pi^{n_{i+1}}(m) - \sum_{m=0}^{n_i} \pi^{n_i}(m)} \\
&> \frac{\mathbb{E}(C(N^{n_i}, S^{n_i}(N^{n_i}))) - \mathbb{E}(C(N^{n_{i-1}}, S^{n_{i-1}}(N^{n_{i-1}})))}{\sum_{m=0}^{n_i} \pi^{n_i}(m) - \sum_{m=0}^{n_{i-1}} \pi^{n_{i-1}}(m)} = W_i.
\end{aligned}
$$

Ohartu $W_{i+1} > W_i$ eta $n_{i+1} > n_i$ indize gaitasuna propietatea inplikatzen dutela.

2.2. Proposizioan aurkeztutako pausuek Whittle indizea definitzen dutela ikusteko hurrengoa erakutsi behar da,

1. -1 atari politika optimoa dela (2.3.3) problemarentzat edozein $W < W_0$-tarako.

2. $n_i < \infty$ atari politika optimoa da (2.3.3) problemarentzat $W_i < W < W_{i+1}$ baldintza betetzen duten W guztientzat.

3. Azkenik ∞ atari politika optimoa da (2.3.3) problemarentzat $\infty > W > W_J$ and $J < \infty$ betetzen duten W guztientzat.

1. pausua frogatzeko, ohartu $W < W_0$ hartuz

$$
\begin{aligned}
&W \sum_{m=0}^{n} \pi^n(m) < \mathbb{E}(C(N^n, S^n(N^n))) - \mathbb{E}(C(N^{-1}, S^{-1}(N^{-1}))), \\
&\Longrightarrow \mathbb{E}(C(N^{-1}, S^{-1}(N^{-1}))) < \mathbb{E}(C(N^n, S^n(N^n))) - W \sum_{m=0}^{n} \pi^n(m), \text{ for all } n \geq 0,
\end{aligned}
$$

hau da, $g^{(-1)}(W) < g^{(n)}(W)$ edozein $n \in \mathbb{N} \cup \{0\}$, eta beraz, $g(W) = g^{(-1)}(W)$. -1 atari politika optimoa da beraz (2.3.3) problemarentzat edozein $W < W_0$.

2. pausua indukzioz frogatuko da. Ohartu n_0-ren definizioa dela eta, edozein $n \geq 0$ hartuz

$$\mathbb{E}(C(N^{n_0}, S^{n_0}(N^{n_0}))) - W_0 \sum_{m=0}^{n_0} \pi^{n_0}(m) \leq \mathbb{E}(C(N^n, S^n(N^n))) - W_0 \sum_{m=0}^{n} \pi^n(m), \tag{2.6.1}$$

betetzen da, hau da, $g^{(n_0)}(W_0) \leq g^{(n)}(W_0)$, edozein $n \geq 0$. Bestalde, $g^{(n_0)}(W_0) \leq g^{(-1)}(W_0)$ berehalakoa da. Hipotesiz $\sum_{m=0}^{n} \pi^n(m)$ hertsiki gorakorra da n-n, eta (2.6.1) ekuaziotik

$$\begin{aligned} &\mathbb{E}(C(N^{n_0}, S^{n_0}(N^{n_0}))) - W \sum_{m=0}^{n_0} \pi^{n_0}(m) \leq \mathbb{E}(C(N^n, S^n(N^n))) - W \sum_{m=0}^{n} \pi^n(m) \\ &\Longrightarrow g^{(n_0)}(W) \leq g^{(n)}(W), \end{aligned}$$

lor daiteke, edozein $n \leq n_0$ eta $W_0 < W$ direnean.

Bereziki, $g^{(n_0)}(W) \leq g^{(n)}(W)$, $W_0 < W < W_1$ betetzen duen edozein W-rako betetzen da non $n \leq n_0$. Modu berean, W_1-en definiziotik $g^{(n_0)}(W_1) \leq g^{(n)}(W_1)$ lor daiteke edozein $n \geq n_0 + 1$, eta gainera $\sum_{m=0}^{n} \pi^n(m)$ hertsiki gorakorra denez $g^{(n_0)}(W) \leq g^{(n)}(W)$ lor daiteke $W_0 < W < W_1$ betetzen duen edozein W-rentzat eta $n \geq n_0 + 1$.

Beraz, $g^{(n_0)}(W) \leq g^{(n)}(W)$ frogatu da edozein n-rako eta $W_0 < W < W_1$, hau da, n_0 atari politika optimoa da W kontsideratuz non $W_0 < W < W_1$. Honek indukzioaren lehenengo pausua ezartzen du $i = 0$. Demagun orain $i - 1 \geq 0$ pausuentzat betetzen dela, hau da, n_i atari politika optimoa dela (2.3.3) problemarentzat, W finko batentzat zeinak $W_{i-1} < W < W_i$ betetzen duen. Demagun $n_i < \infty$. W_i-ren definizioak eta n_{i-1} politika optimoa izateak W finko batentzat non $W_{i-1} < W < W_i$,

$$g^{(n_{i-1})}(W_i) = g^{(n_i)}(W_i) \leq g^{(n)}(W_i), \text{ edozein } n \geq 0,$$

inplikatzen du. Gogoratu $\sum_{m=0}^{n} \pi^n(m)$ hertsiki gorakorra dela n-n eta beraz,

$$g^{(n_i)}(W) \leq g^{(n)}(W), \text{ edozein } n \leq n_i, \text{ edozein } W \text{ non } W_i < W < W_{i+1} \text{ direnean.}$$

Bestalde, W_{i+1}-ren definiziotik

$$g^{(n_i)}(W) \leq g^{(n)}(W), \text{ edozein } n \geq n_i + 1, \text{ eta edozein } W \text{ non } W_i < W < W_{i+1} \text{ direnean,}$$

lor daiteke. Beraz, n_i atari politika (2.3.3) problemarentzat optimoa dela frogatu da W balioak $W_i < W < W_{i+1}$ betetzen badu.

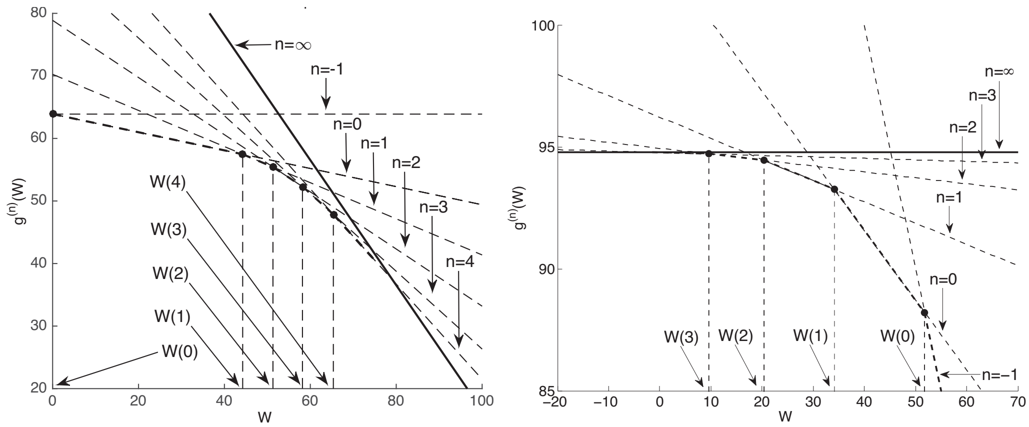
Azkenik, 3. pausua frogatu da $J < \infty$ den kasurako, ohartu goian jarraitu den indukzio argumentuetatik

$$g^{(n_{J-1})}(W_J) = g^{(n_J)}(W_J) \leq g^{(n)}(W_J), \text{ edozein } n \geq 0 \text{ denean,}$$

lortu da, eta $\sum_{m=0}^{n} \pi^n(m)$ hertsiki gorakorra izateak n-n

$$g^{(n_J)}(W) < g^{(n)}(W), n \leq n_J = \infty, \text{ edozein } W \text{ non } W_J < W,$$

inplikatzen du. Honekin teoremaren frogapena amaitzen da.



Irudia 2.3: Behe-inguratzailea irudikatu da, *i.e.,* $g = \min_n\{g^{(n)}\}$. Ezkerraldean: Uzteak gerta daitezkeen eredu klase-anitza non $b^a(m) = \lambda$, $d^a(m) = \mu a + \theta m$, $C(m,a) = (1 + 2\theta)m + 3m^2$, non $a = 0, 1$, eta $\theta = 6, \lambda = 23, \mu = 10$. Eskuinaldean: Inbentario problema bat item galkorrekin non $b^a(m) = \mu a$, $d^a(m) = \lambda + \theta m$, $C(m,a) = (2 + 3\theta)m + m^2 + 14\pi^m(0)$, eta $\mu = 5, \lambda = 10, \theta = 2.5$.

2.6.3 2.1. Korolarioaren frogapena

Notazioa errazteko helburuarekin, klaseekiko dependentzia alde batera utziko da frogapen honetan.

(2.3.3) problemaren soluzio optimo bat atari motako politika denez, batez besteko kostu optimoa n atari politikapean $g(W) := \min_n\{g^{(n)}(W)\}$ izango da W finko batentzat, non 0-1 motako politika bada n

$$g^{(n)}(W) := \sum_{m=0}^{\infty} C(m, S^n(m))\pi^n(m) - W \sum_{m=0}^{n} \pi^n(m), \tag{2.6.2}$$

den eta 1-0 motako politika bada orduan

$$\begin{aligned} g^{(n)}(W) &:= \sum_{m=0}^{\infty} C(m, S^n(m))\pi^n(m) - W \sum_{m=n+1}^{\infty} \pi^n(m) \\ &= \sum_{m=0}^{\infty} C(m, S^n(m))\pi^n(m) + W \sum_{m=0}^{n} \pi^n(m) - W. \end{aligned} \tag{2.6.3}$$

$g(W)$-ren balioa minimizatzen duen politika $n(W)$ izenez deituko da.

$g(W)$, W-rekiko funtzio afin ez-gorakorren behe inguratzailea da. Orduan $g(W)$ funtzio ahurra eta ez-gorakorra da. Ikusi 2.3. Irudia, non $g^{(n)}(W)$-ren balioa irudikatu den n 0-1 motako atari-politika den kasurako ezkerraldean eta 1-0 motako atari-politika den kasurako eskuinaldean.

Berehala ikus daiteke $g(W)$ deribatua eskuinetik W-n $-\sum_{m=0}^{n(W)} \pi^{n(W)}(m)$ dela 0-1 motako atari-politikentzat (eta $\sum_{m=0}^{n(W)} \pi^{n(W)}(m) - 1$ 1-0 atari-politikentzak). $g(W)$ ahurra denez W-n, deribatua eskuinetik ez-gorakorra da W-n. Gainera, $\sum_{m=0}^{n} \pi^n(m)$ hertsiki beherakorra da n-n, eta beraz, $n(W)$ ez-beherakorra (ez-gorakorra) da W-n. Soluzio optimo bat atari motakoa denez, pasiboa izatea optimoa den egoeren multzoa $D(W) = \{m : m \leq n(W)\}$ idatz daiteke 0-1 motako atari-politikentzat (edo $D(W) = \{m : m \geq n(W)\}$ 1-0 motako atari politikentzat). $n(W)$ ez-beherakorra denez (ez-gorakorra), definizioz k *bandit*-ak indize gaitasuna propietatea du.

Izan bedi $\tilde{W}(n)$ batez besteko kostua n eta $n-1$ atari-politikapean berdina izan dadin W subsidioak hartzen duen balioa. Orduan, (2.3.3) erabiliz,

$$\mathbb{E}(C(N^n, S^n(N^n)) - \tilde{W}(n)\mathbb{E}(\mathbb{1}_{S^n(N^n)=0}),$$

eta $\mathbb{E}(C(N^{n-1}, S^{n-1}(N^{n-1})) - \tilde{W}(n)\mathbb{E}(\mathbb{1}_{S^{n-1}(N^{n-1})=0})$ berdinak direla ondoriozta daiteke edozein $n \geq 1$ denean. n atari-politika 0-1 motakoa denean $\mathbb{E}(\mathbb{1}_{S^n(N^n)=0}) = \sum_{m=0}^{n} \pi^n(m)$, orduan $\tilde{W}(n)$ (2.3.6)-k ematen du. n atari-politika 1-0 motakoa denean $\mathbb{E}(\mathbb{1}_{S^n(N^n)=0}) = \sum_{m=n+1}^{\infty} \pi^n(m)$, orduan $\tilde{W}(n)$

$$\begin{aligned}
&\frac{\mathbb{E}(C(N^n, S^n(N^n))) - \mathbb{E}(C(N^{n-1}, S^{n-1}(N^{n-1})))}{\sum_{m=n+1}^{\infty} \pi^n(m) - \sum_{m=n+1}^{\infty} \pi^{n-1}(m)} \\
&= -\frac{\mathbb{E}(C(N^n, S^n(N^n))) - \mathbb{E}(C(N^{n-1}, S^{n-1}(N^{n-1})))}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} = -(2.3.6),
\end{aligned}$$

da.

Froga daiteke $\tilde{W}(n)$ monotonoa izateak $g(\tilde{W}(n)) = g^{(n)}(\tilde{W}(n)) = g^{(n-1)}(\tilde{W}(n))$ inplikatzen duela. Gainera, 0-1 motako atari politikek, $\frac{\mathrm{d}g^{(n)}(W)}{\mathrm{d}W} = -\sum_{m=0}^{n} \pi^n(m)$ n-n beherakorra denez, $g(W) = g^{(n-1)}(W)$ non $\tilde{W}(n-1) \leq W \leq \tilde{W}(n)$ betetzen dute. Honek Whittle indizea $W(n) = \tilde{W}(n)$-k ematea inplikatzen du. Antzeko moduan, 1-0 politiken kasuan, $\frac{\mathrm{d}g^{(n)}(W)}{\mathrm{d}W} = -\sum_{m=n+1}^{\infty} \pi^n(m)$ gorakorra da n-n eta orduan $g(W) = g^{(n)}(W)$ non $\tilde{W}(n) \leq W \leq \tilde{W}(n-1)$. Honek Whittle indizea $W(n) = \tilde{W}(n)$-k ematen duela inplikatzen du.

3. Kapitulua

Indize politika uzteak gerta daitezken ilara batentzat

Contents

3.1 Ereduaren deskribapena eta aurrekariak 34
3.2 Erlaxazioa 36
3.3 Kasu bereziak 39
3.4 M/M/1 ilara klase-anitza 43
3.5 Errendimendu optimoa asintotikoki 45
3.6 Emaitza numerikoak 46
3.7 Eranskina 52

Kapitulu honetan baliabide esleipenerako problema aztertuko da zerbitzari-bakarreko klase-anitzeko ilara batetan, zeinetan matentze kostuak ganbilak diren eta bezeroek sistema utz dezaketen zerbitzua jaso aurretik. Helburua batez besteko kostua minimizatzea izango da. Problema hau 2. Kapituluan aurkeztu den testuinguruan kokatzen da. Bertan aurkeztu den egitura jarraituz, lehenik eta behin optimizazio problema erlaxatuko da, atari politikak optimoak direla frogatuko da problema ez mugatuarentzat, indize gaitasuna betetzen dela ikusiko da eta Whittle indizea lortuko da. Kapituluko gainerako azalpena, Whittle indizearen propietateak garatzean oinarrituko da mugako eremuetan edota sarrerako parametruen aukera partikular batzuetarako. Mantzentze kostuak linealak diren kasuan, Whittle indizea konstantea dela frogatuko da eta ilaran dauden bezero kopuruekiko eta iritsiera tasarekiko askea dela. Mantzentze kostuak ganbilak direnean indizearen propietateak garatu dira mugako eremuetan: indizearen portaera analizatu da hurrengo kasuetan, (i) klase batean zain dauden bezero kopurua hazten denean, indize politikaren egitura asintotikoa garatzea ahalbideratzen du honek, (ii) uzte tasa desagertzen denean, klase-anitzeko M/M/1 ilararako indizeak mantzentze kostuak gangilak diren kasurako berreskuratzea ahalbideratzen du honek, (iii) iritsiera tasa 0-rantz ala ∞-rantz doan heinean, trafiko arinari eta trafiko geldoari dagozkienak, hurrenez hurren. Ez hori bakarrik, Whittle indize politika trafiko arina eta trafiko geldoa den kasuetan optimoa dela frogatu da eta errendimendu ona erakusten duela ikusi da trafiko ezberdinetan.

Kapitulua hurrengo moduan egituratu da. 3.1. Sekzioan uzteak gerta daitezken zerbitzari bakarreko klase-anitzeko ilara eredua deskribatu da. 3.2. Sekzioan erlaxazio Lagrangearra aplikatu da eta atari-politikak optimoak direla frogatu da optimizazio problema ez mugatuarentzat. Indize gaitasuna frogatu da eta Whittle-en indizea kalkulatu da batez besteko kostua minimizatzen den kasurako. 3.3. Sekzioan Whittle indizea kalkulatu da mantentze kostuak linealak diren kasuan eta hainbat mugako eremuetan indizearen propietate garatu dira mantentze kostuak ganbilak direnean. 3.4. Sekzioan M/M/1 ilararentzat indizea kalkulatu da uzteak gerta ezin diren kasuan. 3.5. Sekzioak emaitza asintotikoak aurkezten ditu. Azkenik, 3.6. Sekzioan numerikoki Whittle indize politikaren errendimendua aztertu da. Frogapen gehienak 3.7.1. Eranskinean aurki daitezke.

3.1 Ereduaren deskribapena eta aurrekariak

K bezero klase dituen ilara zerbitzari bakarra kontsideratuko da. k klaseko bezeroak λ_k tasako Poisson prozesu bat jarraituz iristen dira sistemara eta behar duten zerbitzuak banaketa esponentziala jarraitzen du $1/\mu_k$, $k = 1, \ldots, K$. $\rho_k := \lambda_k/\mu_k$ izenez deituko da k klaseari dagokion trafikoaren karga, eta $\rho := \sum_{k=1}^K \rho_k$ sistema guztiko karga. Bezeroen uztean hurrengo moduan definituko dira:

- k klaseko bezero batek ilara uzten du θ_k tasako banaketa esponentziala jarraitzen duen denbora tarte baten ostean, non $k = 1, \ldots, K$, eta $\theta_k > 0$.
- Zerbitzua jasotzen ari den k klaseko bezero batek sistema uzten du θ'_k tasako banaketa esponentziala jarraitzen duen denbora tarte baten ostean, non $k = 1, \ldots, K$, eta $\theta'_k \geq 0$.

Zerbitzariak 1-eko edukiera du eta gehienez bezero bat har dezake zerbitzura uneoro. Bestalde, zerbitzua lehentasunezkoa da, hau da, zerbitzua jasotzen ari den bezeroa baino lehentasun handiagoa duen bezero bat ilaran zain balin badago zerbitzariak lehentasun handiagoa duen bezeroa hartuko du bestea utziz. Hurrengo hipotesia egingo da:

$$\mu_k + \theta'_k \geq \theta_k, \text{ edozein } k\text{-tarako.}$$

Hau da, k klaseko bezeroen irteera tasa handiagoa da bezeroa zerbitzua jasotzen ari denean ilaran zain dagoenean baino.

Uneoro, ϕ politikak erabakitzen du zein klase zerbitzatuko den. Markov propietatea dela bide, erabakiak une horretan dauden klase ezberdinetako bezero kopuruaren araberakoak soilik izango dira. ϕ politika finko bat emanda, $N_k^\phi(t)$ aldagaia t denbora unitatean k klaseko bezero kopuruaren adierazle da, (zerbitzua jasotzen ari den bezeroa barne), eta $\vec{N}^\phi(t) = (N_1^\phi(t), \ldots, N_K^\phi(t))$. Izan bedi $S_k^\phi(\vec{n}) \in \{0, 1\}$ k klaseko bezeroei eskeinitako zerbitzu kapazitatea t denbora unitatean ϕ politikapean eta $\vec{N}(t) = \vec{n}$ egoeran. Zerbitzariaren edukieraren gaineko baldintza $S_k^\phi(\vec{n}) = 0$, $n_k = 0$ bada eta

$$\sum_{k=1}^K S_k^\phi(\vec{n}) \leq 1, \tag{3.1.1}$$

dela kontsideratu da. Bestale, azken baldintza betetzen duten eta onargarriak diren politiken multzoa $\mathcal{U}$ izenpetuko da.

Goian deskribatutako RBP-a jaiotza-eta-heriotza motakoa denak, honako trantsizioak ditu:

$$\vec{n} \to \vec{n} + \vec{e}_k, \lambda_k \text{ tasarekin, eta,}$$

$$\vec{n} \to \vec{n} - \vec{e}_k, \mu_k S_k^\phi(\vec{n}) + \theta_k(n_k - S_k^\phi(\vec{n})) + \theta'_k S_k^\phi(\vec{n}) \text{ tasarekin,}$$

non $n_k > 0$, eta $\vec{e}_k$ K-dimensiotako bektore bat den zeinaren osagarri guztiak 0 diren k. osagarria izan ezik, 1-a baita.

Izan bedi $C_k(n,a)$ denbora unitateko kostua sisteman k klaseko n bezero daudenean eta k klaseko bezero bat zerbitzatua denean ($a = 1$), eta zerbitzurik jasotzen ez duenean ($a = 0$). Demagun $C_k(\cdot, 0)$ eta $C_k(\cdot, 1)$ funtzio ganbilak direla eta ez beherakorrak eta

$$C_k(n,0) - C_k((n-1)^+,0) \leq C_k(n+1,1) - C_k(n,1) \leq C_k(n+1,0) - C_k(n,0), \tag{3.1.2}$$

betetzen dutela edozein $n \geq 0$-rako.

Ohartu $C_k(0,0) \geq C_k(0,1)$ betetzen bada, orduan (3.1.2)-ek $C_k(n,0) \geq C_k(n,1)$ inpikatzen du edozein n-tarako. Ohartu baita (3.1.2) beti betetzen dela (i) $C_k(n,a) = C_k(n)$ denean, edo (ii) $C_k(n,a) = C_k((n-a)^+)$ denean. (i) kasua bezeroak sisteman mantzentzearen kostua adierazten du, ordez, (ii) kasuak bezeroak ilaran mantzentzearen kostua.

Bestalde δ_k kostua definituko da ilara utzi duten k klaseko bezero bakoitzeko, eta δ'_k kostua zerbitzaria utzi duten k klaseko bezero bakoitzeko.

Optimizazioaren helburua OPT politika optimoa aurkitzea da, epe luzerako batez besteko kostuaren itxaropena minimizatzeko, hau da,

$$\mathcal{C}^\phi := \limsup_{T\to\infty} \sum_{k=1}^{K} \frac{1}{T} \mathbb{E}\Bigg(\int_0^T C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t)))\ \mathrm{d}t + \delta_k R_k^\phi(T) + \delta'_k R_k'^\phi(T) \Bigg), \tag{3.1.3}$$

minimizatzen duen ϕ politika aurkitu behar da, non $R_k^\phi(T)$ eta $R_k'^\phi(T)$-k, $[0,T]$ denbora tartean ilara utzi duten eta zerbitzaria utzi duten k klaseko bezero kopurua adierazten duten ϕ politikapean, hurrenez hurren. $\mathcal{C}^{OPT} = \inf_{\phi\in\mathcal{U}} \mathcal{C}^\phi$ da batez besteko kostu optimoaren adierazle.

Hurrengo bi berdinketak,

$$\mathbb{E}(R_k^\phi(T)) = \theta_k \mathbb{E}\left(\int_0^T (N_k^\phi(t) - S_k^\phi(\vec{N}^\phi(t)))\mathrm{d}t \right)$$

eta

$$\mathbb{E}(R_k'^\phi(T)) = \theta'_k \mathbb{E}\left(\int_0^T S_k^\phi(\vec{N}^\phi(t))\mathrm{d}t \right),$$

Dynkin-en formulari esker betetzen dira [3, Chapter 6.5]. Hurrengo notazioa beharrezkoa izango da

$$\tilde{C}_k(n_k,a) := C_k(n_k,a) + \delta_k\theta_k(n_k - a)^+ + \delta'_k\theta'_k \min(a, n_k), a \in \{0,1\}, \tag{3.1.4}$$

non (3.1.3) funtzioa objektiboa

$$\limsup_{T\to\infty} \sum_{k=1}^{K} \frac{1}{T} \mathbb{E}\left(\int_0^T \tilde{C}_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t)))\ \mathrm{d}t \right), \tag{3.1.5}$$

ere idatzi daitekeen.

Goian deskribatu diren kontrol estokastikoko problemak oso konplexuak gerta daitezke ebazterako orduan. Kostuak linealak diren kasu berezian ere, politika optimoen egitura propietatek lortzea erronka latza suposatzen du. Adibidez, [40] artikuluan kontrol dinamiko optimoa aztertu da 2 bezero klaseko ($K = 2$) sistema batean, non $\delta_k = \delta'_k$, $\theta_k = \theta'_k$, $\mu_1 = \mu_2 = 1$, eta kostuak linealak diren, hau da, $C_k(n, a) = c_k n$. Definitu $\tilde{c}_k := c_k + \delta_k \mu_k$. $\tilde{c}_1 \geq \tilde{c}_2$ and $\theta_1 \leq \theta_2$ den kasu berezian, autoreek 1 klaseari lehentasuna ematea dela optimoa frogatu dute, ikusi [40, 3.5. Teorema]. Intuitiboki, argi dago 1 klaseari lehentasuna ematea dela akzio optimoa, 1 klasea zerbitzatzeak miopikoki minimitzatzen baitu mantentze gehi uzte kostua eta gainera, 2 klaseko bezeroak ilaran mantentzea desiragarria da, izan ere uzte tasa handiagoa dute. [24] artikuluan kontrol dinamiko optimoa aztertu da $C_k(n, a) = c_k n$, $\delta_k = \delta'_k$, eta $\theta_k = \theta'_k$ edo $\theta'_k = 0$ kasurako. Klaseak $\tilde{c}_1 \geq \cdots \geq \tilde{c}_K$, $\tilde{c}_1(\mu_1 + \theta'_1 - \theta_1) \geq \cdots \geq \tilde{c}_K(\mu_K + \theta_K - \theta'_K)$ eta $\tilde{c}_1(\mu_1 + \theta'_1 - \theta_1)/\theta_1 \geq \ldots \geq \tilde{c}_K(\mu_K + \theta'_K - \theta_K)/\theta_K$ ordenatuta dauden kasurako, autoreak lehentasuna klaseei $1 > 2 > \ldots > K$ ordenean ematea dela optimoa frogatu dute.

Goian azaldutako parametro berezietatik at, edo mantentze kostuak ganbilak direnean, soluzio optimoa egoeraren menpekoa izatea espero da, eta tesi honen autoreak dakitenez, ez dira egitura emaitzik existitzen kontrol optimo estokastikorako.

Kostuak ganbilak diren kasurako kontrol optimoaren ideia bat lortu ahal izateko, kapitulu honetan optimizazio problemaren bertsio erlaxatu bat ebatzi da, 2. Kapituluan deskribatu den bezela. Azken honek jatorrizko ereduarentzako heuristika bat proposatzea ahalbidetzen du. Heuristika hau trafikoa arina edo trafikoa geldoa denean asintotikoki optimoa dela frogratu da.

3.2 Erlaxazioa

2. Kapituluan aurkeztu den legez, erlaxazio Lagrangearra, problema ergodikoa dela hipotesia eginez, horrela sinplifikatzen da: aurkitu ϕ zeinarentzat

$$\mathcal{C}_k^\phi(W) = \mathbb{E}(\tilde{C}_k(N_k^\phi, S_k^\phi(N_k^\phi))) - W\mathbb{E}(\mathbf{1}_{\{S_k^\phi(N_k^\phi)=0\}}), \tag{3.2.1}$$

minimizatzen den edozein k *bandit*-entzako, non W 2.3. Sekzioan azaldu den legez, pasibo izatearen trukeko subsidio gisa ikus daitekeen. Bestalde, $g_k(W) := \min_\phi \mathcal{C}_k^\phi(W)$ definituko da.

3.2.1. Sekzioan 0-1 motako atari politikak optimoak direla frogatuko da (3.2.1)-rentzat, 3.2.2. Sekzioan klase guztiak indize-gaitasuna dutela frogatuko da eta 3.2.3. Sekzioan Whittle indizea garatuko da zenbait kasu interesgarrietan.

3.2.1 Atari-politikak

Hurrengo proposizioan (3.2.1) problema mugagabearen soluzio optimoa atari-politika bat dela frogatuko da, *i.e.*, bezeroen kopurua n atariaren gainetik dagoenean, klase hori zerbitzatua da, eta azpitik badago ez da zerbitzatuko. $\phi = n$-z deituko dugu n ataritzat duen atari-politika, hau da, $S_k^n(m) = 1$ edozein $m > n$, eta $S_k^n(m) = 0$ bestela.

3.1 Proposizioa. *Izan bedi $n = -1, 0, 1, \ldots,$ orduan existitzen da $\phi = n$ atari-politika (3.2.1) problema mugagabearentzat soluzio optimoa dena.*

Frogapena. k-rekiko menpekotasuna alde batera utzi da frogapen honetan. $V(m)$ balio funtzioak Bellmanen ekuazioa betetzen du batazbesteko kostua kontsideratzen duen ereduarentzat, ikusi 1.3.3. Sekzioa, hau da,

$$\begin{aligned}&(\mu+\theta'+m\theta+\lambda)V(m)+g(W)=\lambda V(m+1)+\theta(m-1)V((m-1)^+)\\&+\min\{\tilde{C}(m,0)-W+(\mu+\theta')V(m)+\theta V((m-1)^+),\tilde{C}(m,1)+(\mu+\theta')V((m-1)^+)+\theta V(m)\},\end{aligned} \tag{3.2.2}$$

non $g(W)$ soluzio optimoan sortzen duen batez besteko kostua den. Atari-politika bat optimoa dela frogatzea, (3.2.2)-n $m+1$, $m\geq 0$ egoeran pasibo izatea optimoa bada orduan (3.2.2) m egoeran ere pasibo izatea optimoa dela frogatzearen baliokidea da, *i.e.*, $\tilde{C}(m+1,0)-W+(\mu+\theta'-\theta)V(m+1)\leq\tilde{C}(m+1,1)+(\mu+\theta'-\theta)V(m)$. Azken honek $\tilde{C}(m,0)-W+(\mu+\theta'-\theta)V(m)\leq\tilde{C}(m,1)+(\mu+\theta'-\theta)V((m-1)^+)$, inplikatzen du. Goiko inekuazioa betetzeko baldintza nahikoak (3.1.2) eta $V(m+1)+V((m-1)^+)\geq 2V(m)$, inekuazioa dira edozein $m\geq 0$-rako. Azken baldintza hau, Balio funtzioa ganbila izatea, behean frogatuko da. Argumentu honek amaitzen du frogapena.

Trantsizio tasak bornatuak direnean, sistema uniformizatu daiteke eta *value iteration* algoritmoa erabil daiteke ganbiltasuna frogatzeko. Halere, tesi honetako trantsizio tasak ez dira bornatuak. SRT metodoa jarraituz 1.3.3. Sekzioan aurkeztu dena, egoeren espazioa moztua kontsidera daiteke, $L>1$ aldagaiarekin moztu dena, eta iritsiera trantsizio tasak doitu hurrengo moduan:

$$q^{\phi,L}(m,m+1):=\lambda\left(1-\frac{m}{L}\right)^+=\lambda\max\left(0,1-\frac{m}{L}\right),$$

edozein $m=0,\dots,L$-rako, linealki txikitzen direnak $q^{\phi,L}(L+1,L)=0$ lortzen den arte. Deitu $V^L(m)$ espazio L-moztuaren balio funtzioari. Zenbait baldintza betetzen direla ziurtatu ondoren, (3.7.1. Eranskinean egin den legez), [25, 3.1. Teorema]-ko emaitzak ziurtatzen du $V^L(m)\to V(m)$ L infinitoruntz doanean. Orduan, V ganbila izatea V^L edozein L-tarako ganbila izatetik ondorioztatzen da, eta beraz azken hau frogatzearekin nahikoa da. Espazio moztua uniformiza daiteke eta beraz, *value iteration* algoritmoa erabil daiteke V^L ganbila dela ikusteko. Frogapena 3.7.1. Eranskinean aurki daiteke. □

Behean egoera egonkorreko banaketa aurkezten da $\phi=n$ atari-politikarentzat. Edozein k-klaserentzat deitu $\pi_k^n(i)$, i egoeran eta n atari-politikapeko egoera-egonkorreko probabilitateari. Orduan

$$\pi_k^n(i)=\prod_{m=1}^{i}\frac{q_k^n(m-1,m)}{q_k^n(m,m-1)}\pi_k^n(0),\quad i=1,2,\dots, \tag{3.2.3}$$

non $\pi_k^n(0)=\left(1+\sum_{i=1}^{\infty}\prod_{m=1}^{i}\frac{q_k^n(m-1,m)}{q_k^n(m,m-1)}\right)^{-1}$ eta

$$\begin{aligned}&q_k^n(m,m-1):=\begin{cases}\theta_k m & \text{edozein } m\leq n \text{ denean},\\ \mu_k+\theta'_k+\theta_k(m-1) & \text{edozein } m>n \text{ denean},\end{cases}\\&q_k^n(m,m+1):=\lambda_k,\qquad \text{edozein } m \text{ denean}.\end{aligned} \tag{3.2.4}$$

3.1 Oharra. *3.1. Proposizioan (3.2.1)-rentzat atari-politikak optimoak direla ikusi da $\mu_k + \theta'_k \geq \theta_k$ den kasuan eta (3.1.2) betetzen denean. Ordez, $\mu_k + \theta'_k < \theta_k$ balitz, eta $\tilde{C}_k(m, 1) > \tilde{C}_k(m, 0)$ edozein m-tarako (kasu honetan ez da (3.1.2) betetzea eskatzen), orduan ($W \geq 0$ bada) politika optimoa m egoera guztietan pasibo izatea da. Hau berehalakoa da* (3.2.2) *ekuaziotik, k azpi-indizea gehituz, izan ere, beti pasibo izatea optimoa da baldin edozein m-rako*

$$\tilde{C}_k(m, 0) - W + (\mu_k + \theta'_k - \theta_k)V_k(m) \leq \tilde{C}_k(m, 1) + (\mu_k + \theta'_k - \theta_k)V_k((m-1)^+).$$

Azken hau goiko hipotesien ondorioa da eta $V_k(\cdot)$ balio funtzioa ez-beherakorra izatearena. $V_k(\cdot)$ funtzio ez-beherakorra izatearen frogapena 3.7.1. Eranskinetik ondorioztatzen da.

Beste kasuetan, numerikoki ikusi da atari-politikak optimoak direla, hala ere, ez izan da matematikoki ezarri hori egia denik.

3.2.2 Indize-gaitasuna

Sekzio honetan kapitulu honetan aztergai den ereduarentzat, bezero klaseek inideze-gaitasuna dutela erakutsiko da.

3.2 Proposizioa. *Klaseek indize-gaitasuna dute.*

Frogapena. Frogapen honetan k-rekiko dependentzia alde batera utzi da.

2.2. Proposizioaren frogapenetik, indize gaitasuna $\sum_{i=0}^{n} \pi^n(i)$ hertsiki gorakorra izateak inplikatzen duela ikusi da. Horregatik $\sum_{i=0}^{n} \pi^n(i)$ hertsiki gorakorra dela ikustea nahikoa da, edo baliokideki, $1 - \sum_{i=n+1}^{\infty} \pi^n(i)$ hertsiki beherakorra dela. (3.2.3) erabiliz, azken hau frogatzeko nahikoa da

$$\frac{\sum_{m=n+1}^{\infty} \prod_{i=1}^{m} \frac{q^n(i-1,i)}{q^n(i,i-1)}}{\sum_{m=n}^{\infty} \prod_{i=1}^{m} \frac{q^{n-1}(i-1,i)}{q^{n-1}(i,i-1)}} < \frac{1 + \sum_{m=1}^{\infty} \prod_{i=1}^{m} \frac{q^n(i-1,i)}{q^n(i,i-1)}}{1 + \sum_{m=1}^{\infty} \prod_{i=1}^{m} \frac{q^{n-1}(i-1,i)}{q^{n-1}(i,i-1)}}, \tag{3.2.5}$$

betetzen dela ikustea, edozein n-tarako, non $q^n(\cdot, \cdot)$ trantsizioak (3.2.4)-n definitu diren. Ohartu $q^n(m-1, m) = q^{n-1}(m-1, m)$ edozein m-rako eta $q^n(m, m-1) = q^{n-1}(m, m-1)$ edozein $m \neq n$-rako. $\mu + \theta' \geq \theta$ denaren hipotesitik $q^n(n, n-1) \leq q^{n-1}(n, n-1)$ lortzen da. Beraz, (3.2.5) ekuazioko ezkerraldeko espresioa 1 baino txikiagoa da, eskuinaldeko espresioa 1 baino handiagoa den heinean. Honek (3.2.5) inplikatzen du. □

3.2.3 Whittle indizea

Kapitulu honetan kontsideratu den eredurako 0-1 atari-politikak optimoak direla eta indize gaitasuna duela ikusi ostean, Whittle indizea 2.2. Proposizioan aurkeztu bezala kalkula daiteke. (2.3.6) indizea n-n ez-beherakorra den kasuan Whittle indizea (2.3.6)-k definitzen du.

Ez da posible izan Whittle indizea $W_k(n)$, (2.3.6) ekuazioan emanda bezela ez-beherakorra denik frogatzea orokorrean. Halere, hainbat kasu partikularretan propietate hori froga daiteke. Adibidez,

- $\mu_k + \theta'_k = \theta_k$ den kasuan, $\pi_k^n(m) = \pi_k^{n-1}(m)$ lortzen da edozein m-rako. Beraz, (2.3.6) hurrengo moduan idatz daiteke:

$$\frac{\tilde{C}_k(n, 0)\pi_k^n(n) - \tilde{C}_k(n, 1)\pi_k^{n-1}(n)}{\pi^n(n)} = \tilde{C}_k(n, 0) - \tilde{C}_k(n, 1). \tag{3.2.6}$$

(3.1.2) baldintzak $\tilde{C}_k(n,0) - \tilde{C}_k(n,1) \leq \tilde{C}_k(n+1,0) - \tilde{C}_k(n+1,1)$ inplikatzen du eta beraz, (2.3.6) ez-beherakorra da n-n. Whittle indizea ondorioz (3.2.6) ekuazioak definitzen du.

- 3.3. Proposizioan $C_k(n,a)$ n-n lineala den kasuan, (2.3.6) konstantea dela frogatuko da eta beraz ez-beherakorra n-n. Orduan, Whittle indizea (2.3.6) ekuazioak definitzen du.
- $\theta_k \to 0$ hartuz 3.7. Proposizioan

$$\lim_{\theta_k \to 0} \theta_k \frac{\mathbb{E}(\tilde{C}_k(N_k^n, S_k^n(N_k^n))) - \mathbb{E}(\tilde{C}_k(N_k^{n-1}, S_k^{n-1}(N_k^{n-1})))}{\sum_{m=0}^{n} \pi_k^n(m) - \sum_{m=0}^{n-1} \pi_k^{n-1}(m)},$$

n-n ez-beherakorra dela lortzen da.

Orain hurrengoa frogatuko da $\tilde{C}_k(m_k,0) \geq \tilde{C}_k(m_k,1)$ edozein m_k-rako betetzen denean, Whittle indizea $W_k(n_k)$ beti positiboa da. Hau ikusi ahal izateko, gogoratu $W_k(n_k)$ n_k atari-politika problema mugagabearen soluzioa izan dadin behar den W-ren baliorik txikieran dela. Beraz, edozein $m_k \leq n_k$-rako, optimoa da klasea pasibo mantentzea, hau da, $\tilde{C}_k(m_k,0) - W_k(n_k) + (\mu_k + \theta_k' - \theta_k)V(m_k) \leq \tilde{C}_k(m_k,1) + (\mu_k + \theta_k' - \theta_k)V(m_k - 1)$, 3.1. Proposizioan ikusi den legez. Bestalde, $\tilde{C}_k(m_k,0) \geq \tilde{C}_k(m_k,1)$ denez, $\mu_k + \theta_k' \geq \theta_k$ eta $V(\cdot)$ ez-beherakorra denez (ikusi 3.1. Proposizioaren frogapena), $W_k(n_k) \geq 0$ inpikatzen dute.

Ordez, $\tilde{C}(m_k,0) < \tilde{C}(m_k,1)$ betetzen denean edozein m_k-rako, $W_k(n_k)$ negatiboa izan daiteke hainbat n_k egoerentzat. Adibidez, $\theta_k' = \theta_k$ eta $\delta_k' \gg \delta_k$ direnean. Orduan, nahiz eta k klasearen irteera tasa k klasea zerbitzatzen denean den altuena ($\mu_k + \theta_k' \geq \theta_k$), hainbat n_k egoeratan hobe gerta daiteke k klasea ez zerbitzatzea. Azken hau ulertzeko ohartu k kaseko bezero batek zerbitzua jasotzen ari den bitartean uzten badu sistema, eragindako kostua altuagoa da ilaran zain dagoenean uzten badu sistema baino. Beraz, subsidio negatibo bat, hau da, kostu bat, ordaindu behar da k klasea zerbitzatzea optimoa izan dadin.

Ikuspuntu praktikotik, Whittle indizearen $W_k(n_k)$-ren interesa, 2.2. Proposizioan definitu den bezala, ez duela beste klaseekiko dependentzian datza, hau da, k klaseari dagokion indizea beste $j \neq k$ klaseetan dauden bezero kopuruarekiko askea dela. Beraz, erabilgarriak diren eta errendimendu ona erakusten duten politikak lortzeko modu bat eskeintzen du Whittle indizeak, ikusi 3.6. Sekzioa, eta gaineara politika hau asintotikoki optimoa da hainbat eremuetan, ikusi 3.5. Sekzioa.

3.3 Kasu bereziak

Sekzio honetan 2.2. Sekzioan lortu den Whittle indizearen propietateak sakonago aztertuko dira. 3.3.1. Sekzioan mantentze kostuak linealak diren kasua aztertuko da eta indizea egoeraren menpekoa dela ikusiko da. 3.3.2. Sekzioan mantentze kostu ganbil orokorrentzat propietate asintotikoak garatuko dira.

3.3.1 Mantentze-kostu linealak

Hemen mantentze-kostu linealak dira aztergai, hau da, $C_k(n_k,a) = c_k(n_k - a)^+ + c_k' \min(n_k,a)$. Beraz, funtzio hauek kontsideratuz gero, edozein k klasetako bezeroak ilarari c_k kostua eragiten diote, eta zerbitzua jasotzen ari den k klaseko bezeroek c_k' kostua eragiten dute. Bereziki, $c_k' = c_k$ bada, orduan C_k bezeroek sistemari eragiten dioten kostua lineala da, eta $c_k' = 0$ bada, orduan C_k ilaran zain dauden bezeroen kostua lineala da. Bi kostu funtzio hauek uzteak kontsideratu dituzten hainbat ikerketa artikulutan izan dira aztergai, adibidez [15] artikuluan lehenengoa kontsideratu da, eta [8] artikuluan azkenengoa. (2.3.6)

$W_k(n_k)$	$\theta'_k = \theta_k, \delta'_k = \delta_k$	$\theta'_k = 0$
$c'_k = c_k$	$\frac{\tilde{c}_k \mu_k}{\theta_k}$	$\frac{\tilde{c}_k \mu_k}{\theta_k} - c_k$
$c'_k = 0$	$\frac{\tilde{c}_k \mu_k}{\theta_k} + c_k$	$\frac{\tilde{c}_k \mu_k}{\theta_k}$

Taula 3.1: 3.3. Proposizioan aurkeztutako $W_k(n_k)$-ren balioa, mantentze kostuak linealak direnean.

formulatik Whittle indizearen karakterizazio osoa lor daiteke. Interesgarria da baita, Whittle indizea egoerarekiko askea dela eta baita λ_k iritsiera tasarekiko ere.

Hurrengo, $\tilde{c}_k := c_k + \delta_k \theta_k$, $k = 1, \ldots, K$ definituko da, denbora unitateko ilaran dauden k bezeroen klaseak eragiten duten kostua bezala uler daitekeena. Era berean, $\tilde{c}'_k := \tilde{c}'_k + \delta'_k \theta'_k$ definituko da, denbora unitateko zerbitzua jasotzen ari diren k klaseko bezeroen kostua bezala uler daitekeena.

Orain mantentze kostuak linealak diren kasurako Whittle indizea kalkulatuko da. Frogapena 3.7.2. Eranskinean aurki daiteke.

3.3 Proposizioa. *Izan bedi* $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k \min(n_k, a)$. *Orduan, k klaseari dagokion Whittle indizea*

$$W_k(n_k) = \frac{\tilde{c}_k(\mu_k + \theta'_k)}{\theta_k} - \tilde{c}'_k, \tag{3.3.1}$$

da, edozein n_k*-rako.*

(3.3.1) indizearen ezaugarri interesgarri da λ_k iritsiera tasarekiko eta k klaseko bezero kopuruarekiko, n_k-rekiko askea dela da. 3.3.2. Sekzioan ohar honek mantentze kostuak linealak diren kasurako bakarrik balio duela ikusiko da.

(3.3.1) indizeak hurrengo interpretazioa du. Demagun k klaseko bezero bakarra dagoela sisteman eta ez direla iritsiera gehiago gertatuko, orduan $\tilde{C}_k(1,1) = \tilde{c}'_k$, $\tilde{C}(1,0) = \tilde{c}_k$, $q^1_k(1,0) = \theta_k$, $q^0_k(1,0) = \mu_k + \theta'_k$. (3.3.1) indizea $(\mu_k + \theta'_k)\left(\frac{\tilde{c}_k}{\theta_k} - \frac{\tilde{c}'_k}{\mu_k + \theta'_k}\right)$ bezala idatz daiteke baliokideki, azken hau

$$q^0_k(1,0)\left(\frac{\tilde{C}(1,0)}{q^1_k(1,0)} - \frac{\tilde{C}_k(1,1)}{q^0_k(1,0)}\right), \tag{3.3.2}$$

bezala ere idatz daiteke. Beraz, indizea hurrengo moduan ere interpreta daiteke, k bandit-a aktibo egitearen kostuaren eta pasibo uzteak eragindako kostuaren arteko diferentzia da aktiboa deneko irteera denbora tarte batean.

Orain beste hainbat artikulutan aztergai izan diren adibide batzuk aztertuko dira, ikusi 3.1. Taula ere. Adibidez, lehenik eta behin bezero guztiek sistema utz dezakeleta kontsideratuko da, *i.e.*, $\theta'_k = \theta_k$, edozein $k = 1, \ldots, K$ denean, eta sistema uztearen kostua berdina dela akzio aktibo eta pasibopean, beraz, $\delta_k = \delta'_k$. Sisteman dauden bezero guztiak mantentze kostu bat eragiten dute. Honek $c_k = c'_k$ bultzatzen du eta beraz, $\tilde{c}_k = \tilde{c}'_k$. Hau, (3.3.1) ekuazioan ordezkatuz, $W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k}$ lortzen da. Ilaran zain dauden bezeroak direnean mantentze kostua eragiten duten bezero bakarrak, *i.e.*, $c'_k = 0$ denean, $\tilde{c}_k - \tilde{c}'_k = c_k$ lortzen da eta (3.3.1) ekuazioan ordezkatuz $W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k} + c_k$ lortzen da.

Orain, ilaran dauden bezeroak bakarrik utzi dezaketela sistema onartuko da, hau da, zerbitzua jasotzen ari diren bezeroak ezin dute sistema utzi, beraz, $\theta'_k = 0$, edozein $k = 1, \ldots, K$ denean. Hau da [8] eta [15] artikuluetan kontsideratu den kasua. Lehenik eta behin, mantentze kostua sisteman dauden bezero guztiek eragiten dutela onartuko da, hau da, $c_k = c'_k$ eta beraz, $\tilde{c}'_k = c_k$. (3.3.1) ekuaziotik $W_k(n_k) = \frac{\tilde{c}_k \mu_k}{\theta_k} - c_k$

lortzen da. Era berean, mantentze kostua ilaran dauden bezeroek soilik eragiten duten kasuan, *i.e.*, $c'_k = 0$, $W_k(n_k) = \frac{\tilde{c}_k\mu_k}{\theta_k}$ lortzen da. Azken bi indize hauek [8] eta [15] artikuluetan lortu dira, hurrenez hurren. Bereziki, [15] artikuluan $\frac{\tilde{c}_k\mu_k}{\theta_k} - c_k$ indizea lortu da bezero bakarreko sistema bat aztertuz iritsierarik gertatzen ez den kasuan. Interesgarria da iritsierak kontsideratuz ere indize berdina berreskuratzen dela. Zerbitzua jasotzen ari den bezeroak mantentze kosturik eragiten ez duenean, kapitulu honetan deskribatutako eredua [8]-n kontsideratu den ereduaren baliokidea da, non $\frac{\tilde{c}_k\mu_k}{\theta_k}$ erregela asintotikoki fluido optimoa dela ikusi den zerbitzari anitzeko ilara batean eta trafikoa geldoa denean ($\rho > 1$). Beraz, tesi honetan garatu den Whittle indizeak literaturan lortu diren beste hainbat indize berreskuratzen ditu, sistema parametro berezientzako aztertzen denean.

Sekzio honekin amaitzeko 3.3. Proposizioan lortu den soluzioa ulertzeko intuizioa eskeiniko da $\theta'_k = \theta_k$ eta $c_k = c'_k$ diren kasuan. Kontextu honetan, edozein denbora unitatetan, sisteman dauden bezero guztiek c_k mantentze kostua eragiten dute eta sistema utz dezakete θ_k tasa batekin. $\mathbb{E}(\tilde{C}_k(N_k^{n_k}, S_k^{n_k}(N_k^{n_k}))) = \tilde{c}_k\mathbb{E}(N_k^{n_k})$ espresioa eta $W_k(n_k) = \frac{\tilde{c}_k\mu_k}{\theta_k}$ espresioa (2.3.6) indizean ordezkatuz

$$\theta_k(\mathbb{E}(N_k^{n_k-1}) - \mathbb{E}(N_k^{n_k})) = \mu_k\left(\sum_{m=n_k}^{\infty}\pi_k^{n_k-1}(m) - \sum_{m=n_k+1}^{\infty}\pi_k^{n_k}(m)\right),$$

lortzen da, eta azken hau tasen kontserbazio gisa ikus daiteke. Ez hori bakarrik, ekuazioaren ezkerraldean dagoen gaia n_k eta $n_k - 1$ politikak konparatzean lortzen den denbora unitate batean sistema utzi duten batez besteko bezeroen diferentzia da. Ekuazioaren eskuinaldeko gaia n_k eta $n_k - 1$ politikak konparatzean lortzen den eta denbora unitateko zerbitzua jasotzen ari diren batez besteko bezeroen arteko diferentzia da. Ezkerraldea eta eskuinaldea berdinak izateak tasen kontserbazioa adierazten du.

3.3.2 Mantentze kostu ganbilak

Sekzio honetan Whittle-en indizea karakterizatuko da, $W_k(n_k)$ indizea (2.3.6) Ekuazioak ematen duenean, mantentze kostu funtzioak ez-beherakor eta ganbilak direnean.

Ohartu lehenik eta behin bezeroen uzteei dagokien kostuak funtzio linealak direla. Beraz, 3.3. Proposizioko emaitza erabil daiteke Whittle indizea hurrengo erara idazteko:

$$W_k(n_k) = \delta_k(\mu_k + \theta'_k) - \delta'_k\theta'_k + W_k^c(n_k), \tag{3.3.3}$$

non

$$W_k^c(n_k) := \frac{\mathbb{E}(C_k(N_k^{n_k}, S^{n_k}(N_k^{n_k}))) - \mathbb{E}(C_k(N_k^{n_k-1}, S^{n_k-1}(N_k^{n_k-1})))}{\sum_{m=0}^{n_k}\pi_k^{n_k}(m) - \sum_{m=0}^{n_k-1}\pi_k^{n_k-1}(m)}$$

mantentze kostuari dagokion gaia den. Sekzio honen gainerakoan $W_k^c(n_k)$ izango da aztergai.

3.3.2. Sekzioan egoeraren balio altuentzako Whittle indizea karakterizatu da. 3.3.2. eta 3.3.2. Sekzioetan Whittle indizea lortu da $\lambda_k \downarrow 0$ eta $\lambda_k \uparrow \infty$ kasuetan, trafiko arina eta trafiko geldoaren adierazle direnak, hurrenez hurren. Kasu guztietan, Whittle indizea mantentze funtzioa ez-lineala denean n_k egoeraren menpekoa dela ikusi da.

Whittle indizea egoera handientzat

Sekzio honetan $C_k(n_k, 1)$ eta $C_k(n_k, 0)$ mantentze kostuak maila finituko polinomioek bornatzen dituztela onartuko da, hau da, $P_k < \infty$ eta $Q_k < \infty$ mailek, hurrenez hurren. Beraz, $C_k(n_k, a) = E_k(n_k, a) + o(1)$ idatz daiteke, n_k-ren balio handientzat, non $E_k(n_k, 1) = \sum_{i=0}^{P_k} C_k^{(P_k,i)} n_k^i$ eta

$$C_k^{(P_k,i)} := \lim_{n_k \to \infty} \frac{C_k(n_k, 1) - \sum_{j=i+1}^{P_k} C_k^{(P_k,j)} n_k^j}{n_k^i},$$

eta $E_k(n_k, 0) = \sum_{i=0}^{Q_k} E_k^{(Q_k,i)} n_k^i$, non

$$E_k^{(Q_k,i)} := \lim_{n_k \to \infty} \frac{C_k(n_k, 0) - \sum_{j=i+1}^{Q_k} E_k^{(Q_k,j)} n_k^j}{n_k^i}.$$

Orokortasunik galdu gabe P_k-k $C_k^{(P_k,P_k)} > 0$ eta Q_k-k $E_k^{(Q_k,Q_k)} > 0$ betetzen dutela onartu da.

Hurrengo proposizioan Whittle indiziaren espresioa aurkeztuko da egoera handientzat. Frogapena 3.7.3. Eranskinean aurki daiteke.

3.4 Proposizioa. *Demagun Whittle indizea* (2.3.6) *ekuazioak definitzen duela. Izan bedi* $C_k(n_k, 1)$ *eta* $C_k(n_k, 0)$ P_k *eta* Q_k *mailako polinomioez bornatutako kostu funtzio bi. Orduan,* $W_k(n_k) = W_k^\infty(n_k) + o(1)$ *betetzen da, non* $W_k^\infty(n_k) := \delta_k(\mu_k + \theta_k') - \delta_k'\theta_k' + \tilde{W}_k^c(n_k)$ *eta*

$$\begin{aligned}\tilde{W}_k^c(n_k) :=& (E_k(n_k, 0) - E_k(n_k, 1)) + (\mu_k + \theta_k' - \theta_k)/\theta_k \\ &\cdot \left(\sum_{i=1}^{Q_k} E_k^{(Q_k,i)} n_k^{i-1} + \sum_{i=2}^{P_k} C_k^{(P_k,i)} \sum_{j=0}^{i-2} n_k^{i-2-j} \left(\frac{\lambda_k}{\theta_k}\right)^{j+1} \right).\end{aligned} \tag{3.3.4}$$

$W_k^\infty(n_k)$ *indizea funtzio ez-beherakorra da.*

Demagun $C_k(n_k, a) = C_k(n_k)$ edo $C_k(n_k, a) = C_k((n_k - a)^+)$ non $P_k \geq 2$. Kasu horretan, $P_k = Q_k$ eta $C_k^{(P_k,P_k)} = E_k^{(Q_k,Q_k)}$. Egoera nahiko handia denean, $W_k^\infty(n_k)$-ren balioa maila altueneko polinomioak erabakitzen du, eta

$$\left(E_k^{(P_k,P_k-1)} - C_k^{(P_k,P_k-1)} + \frac{\mu_k + \theta_k' - \theta_k}{\theta_k} E_k^{(P_k,P_k)} \right) n_k^{P_k-1}, \tag{3.3.5}$$

gaiak definitzen du hori. Azken hau, λ_k iritsiera tasarekiko askea da eta beraz, W_k^∞ egoera altuetarako ere iritsiera tasarekiko askea da. (3.3.5) indizea sendoa da eta Whittle indizea hurbiltzeko erabil daiteke sisteman bezero kopurua handia denean. 3.6. Sekzioan numerikoki $W^\infty(\cdot)$ indize politikaren errendimendua aztertu da.

Indizea trafikoa arinean

Hurrengo proposizioan Whittle indiziaren espresioa aurkeztuko da $\lambda_k \downarrow 0$ den kasuan, trafiko arina bezala ezagutzen den eremuan. Frogapena 3.7.4. Eranskinean aurki daiteke. Trafikoa arina dela onartuz, indizea esplizituki aurki daiteke. 3.5. Sekzioan espresio hau erabiliko da Whittle indizea asintotikoki optimoa dela ikusteko trafiko arinean.

3.5 Proposizioa. *Demagun* $W_k(n_k)$ *Whittle indizea* (2.3.6) *ekuazioak definitzen duela. Orduan,* $W_k(n_k) = \delta_k(\mu_k + \theta'_k) - \delta'_k\theta'_k + W^c_k(n_k)$*, non*

$$\lim_{\lambda_k \downarrow 0} W^c_k(n_k) = C_k(n_k, 0) - C_k(n_k, 1) + (C_k(n_k, 0) - C_k(0, 0))\frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k n_k}.$$

$C_k(0, 0) = 0$ hipotesipean, goiko indizea hurrengo moduan idatz daiteke:

$$\lim_{\lambda_k \downarrow 0} W^c_k(n_k) = (\mu_k + \theta'_k + \theta_k(n_k - 1))\left(\frac{C_k(n_k, 0)}{\theta_k n_k} - \frac{C_k(n_k, 1)}{\mu_k + \theta'_k + \theta_k(n_k - 1)}\right).$$

Honek indizearen hurrengo interpretazioa eskeintzen du trafiko arinerako. k klaseko bezero baten egoera n_k bada, eta ez bada iritsierarik gertatuko etorkizunean, indizeak k *bandit*-a aktibo egitearen eta pasibo uztearen arteko kostu ezberdintasuna neurtzen du, aktibo den fase batean irteera batek behar duen denbora tarte baterako.

Indizea trafikoa geldoa denean

Hurrengo proposizioan Whittle indizea aurkeztuko da $\lambda_k \uparrow \infty$, trafiko geldoko eremua bezela ezagutzen dena. Frogapena 3.7.5. Eranskinean aurki daiteke. Trafikoa geldoa dela onartuz, indizea esplizituki aurki daiteke.

3.6 Proposizioa. *Demagun* $W_k(n_k)$ *Whittle indizea* (2.3.6) *ekuazioak definitzen duela. Eta definitu*

$$W^{HT}_k(n_k) := C_k(n_k, 0) - C_k(n_k, 1) + \frac{\mu_k + \theta'_k - \theta_k}{\theta_k}\frac{\mathbb{E}(C_k(N^{n_k-1}_k, 1))}{\lambda_k/\theta_k},$$

non $N^{n_k-1}_k$ *k-klaseko bezero konpuruaren oreka egoeraren adierazle da* $n_k - 1$ *politikapean, eta* (3.2.4) *trantsizio tasek definitzen dute.* $z \geq 1$ *existitzen bada zeinak* $\frac{\mathbb{E}(C_k(N^{n_k-1}_k, 1))}{\lambda^z_k} \to 0$ *betetzen duen,* $\lambda_k \to \infty$*-rantz doanean, orduan,* $W_k(n_k) = \delta_k(\mu_k + \theta'_k) - \delta'_k\theta'_k + W^{HT}_k(n_k) + o(1)$ $\lambda_k \to \infty$*-rantz doanean.*

3.4 M/M/1 ilara klase-anitza

M/M/1 ilara klase anitza oso aztertua izan da hainbat ikerketa arlotan. Mantentze kostuak linealak diren kasuan $c\mu$ indize politika optimoa dela frogatu da bi kasutan: (i) zerbitzu denborak banaketa esponentziala jarraitzen dutenean eta *scheduling*-a lehentasunezkoa denean, ikusi [33], eta (ii) zerbitzu denborak banaketa orokorra dutenean eta *scheduling*-a lehentasunezkoa ez denean, ikusi [45]. Indize politika bat optimoa izatea hurrengo moduan azal daiteke, k klaseari dagokion denbora unitateko mantentze kostu linear bat c_k kostuarekin, zerbitzu bat betetzean c_k irabazi bat jasotzearen problemaren baliokidea da (mantentze kosturik eragin gabe) [47, Section 4.9]. Azken hau MABP bat bezala ikus daiteke, zeinentzat indize politika bat (kasu honetan $c\mu$) optimoa den[1]. Halere, baliokidetza hau mantzentze kostu linealentzat bakarrik da zuzena, honek azaltzen du zergatik kostu orokorrentzat soluzio optimoaren egitura ez den indize motakoa. Kontextu honetan, hurbilketa matematiko emankor bat errendimendu ona erakusten duten indize politikak garatzea izan da, asintotikoki optimoak direla froga daitezkeenak mugako eremuetan, ikusi 1. Kapituluan egindako aipamenak.

[1]Hau MABP-ren *tax* formulazio gisa ikus daiteke, ikusi [47, Section 4.9].

Sekzio honetan indize politika bat garatu da M/M/1 ilara klase-anitzarentzat tesi honetan garatu den Whittle indizearen limitea hartuz uzte tasa 0-rantz doan heinean. Ohartu $W_k(n_k)$ indizea ∞-rantz doala $\theta_k \to 0$-rantz doan heinean, eta indizea θ_k-rekin doituz gero, ez tribiala den limite bat lortzen da. Hurrengo proposizioaren frogapena 3.7.6. Eranskinean aurki daiteke.

3.7 Proposizioa. *Demagun $C_k(n_k, a) = C_k(n_k)$, $a = 0, 1$, $\theta'_k = \theta_k$, eta $\delta_k = \delta'_k = 0$. Orduan,*

$$\lim_{\theta_k \to 0} \theta_k W_k(n_k) = \frac{\mu_k(1-\rho_k)}{\rho_k} \cdot \left[\sum_{m=0}^{\infty} \rho_k^m (1-\rho_k) C_k(n_k - 1 + m) - C_k(n_k - 1) \right]. \tag{3.4.1}$$

Ohartu $C(\cdot)$ funtzioa ganbila izateak, (3.4.1) funtzio ez-beherakorra izatea inplikatzen du.

Mantentze kostua minimizatzea helburutzat duen M/M/1 ilararentzat heuristika bat hurrengo moduan garatu daiteke. Finkatu $\theta_k = \theta'_k$ edozein k-tarako eta kontsideratu indizea θ_k-z biderkatuta $\theta_k \to 0$ den heinean. Heuristika bat lehentasuna (3.4.1) indizearen arabera ematea litzateke.

Orain M/M/1 ilararentzat metodo zuzen bat erabiliz indize politikak garatzea zergatik den ezinezkoa azalduko da. Hartu M/M/1 ilara bat 0-1 egitura duen n politikapean. Sistema hau FIFO zerbitzu mekanismoa jarraitzen duen M/M/1 ilararen baliokidea da, ilaran uneoro n bezero baleude onartuz. Jakina da prozesua orekan dagoenean n egoeran egoteko probabilitatea $1-\rho_k$ dela. Beraz, $\sum_{m=0}^{n} \pi_k^n(m) = 1-\rho_k$ edozein n-rako. Batez beste lortutako subsidioa beraz $W(1-\rho_k)$ da, n politikarekiko askea dena. Honek, subsidioak atari politiken arteko ezberdinketa egitea ezinezkoa izatea inplikatzen du. Arazo honi aurre egiteko, 3.7. Proposizioan bezeroen uzteak gerta daitezkeela onartuko da eta Whittle indizea kalkulatuko da $\theta_k \to 0$ den heinean. [47, 6.5. Sekzioa] ikerketa artikuluan indize berbera lortu da. Lehenik eta behin Whittle indizea garatzen dute artikulu horretako autoreek mantentze kostu deskontaturako. Kasu honetan ere, indize gaitasuna frogatu behar da. Gero, deskontu faktorea 0-runtz hartuz 3.7. Proposizioko indize politika lortzen dute.

n_k egoeraren balio handietarako, (3.4.1) indizea $C'_k(n_k)\mu_k$ funtziotik gertu dago, azken indize politika honi $Gc\mu$ politika deituko zaio. Politika hau [69] artikuluan proposatu zen atzerapen kostu ganbilentzat. $Gc\mu$ politikarekiko baliokidetasuna hurrengo erara ikus daiteke. Demagun n_k handia dela,

$$\begin{aligned}
&\sum_{m=0}^{\infty} \rho_k^m (1-\rho_k) C_k(n_k - 1 + m) - C_k(n_k - 1) \sum_{m=0}^{\infty} \rho_k^m (1-\rho_k) \\
&= (1-\rho_k) \sum_{m=0}^{\infty} \rho_k^m (C(n_k - 1 + m) - C(n_k - 1)) \approx (1-\rho_k) \sum_{m=0}^{\infty} m \rho_k^m C'(n_k - 1) = C'(n_k) \frac{\rho_k}{(1-\rho_k)},
\end{aligned}$$

non n_k handia denean m-rekiko, $\frac{C(n_k-1+m)-C(n_k-1)}{m} \approx C'(n_k)$ lortzen den eta m-ren balioa handia denean, batukarian ez du eraginik. Beraz, (3.4.1)-tik $\lim_{\theta_k \to 0} \theta_k W_k(n_k) \approx C'_k(n_k)\mu_k$ lor daiteke.

Adibide numerikoa. 6.2. Taulan $C'(n)\mu$ eta (3.4.1) indize politiken errendimendua erkatu da uzterik gabeko M/M/1 ilara baterako. $\theta_k = 0$ denan edozein k-tarako, $\sum_{k=1}^{K} \rho_k < 1$ hipotesia behar da sistema egonkorra izan dadin. Demagun 4 bezero klase daudela sisteman eta $\mu_1 = 16, \mu_2 = 27, \mu_3 = 12$, $\mu_4 = 21$, $\rho_1 = 3\rho/9, \rho_2 = \rho/9$, $\rho_3 = 5\rho/9$ eta $\rho_4 = \rho/9$. Mantentze kostuak klase bakoitzerako kubikoak dira, $C_k(n_k) := \alpha_k + \beta_k n_k + \gamma_k n_k^2 + \delta_k n_k^3$, non (3.4.1) ekuazioak:

$$\beta_k \mu_k + \gamma_k \mu_k \left(\frac{3\rho_k - 1}{1-\rho_k} + 2n_k \right) + \delta_k \mu_k \left(3n_k^2 + 3\left(\frac{2\rho_k - 1}{1-\rho_k} \right) n_k + \frac{4\rho_k^2 + \rho_k + 1}{(1-\rho_k)^2} \right),$$

ρ	0.11	0.21	0.31	0.41
(3.4.1)	4.25e-06	1.51e-05	6.07e-06	5.02e-07
$C'(n)\mu$	0.0072	0.0636	0.1002	0.1320
ρ	0.51	0.61	0.71	0.81
(3.4.1)	0.008	0.0291	0.0919	1.7129
$C'(n)\mu$	0.1689	0.3616	1.8280	4.9539

Taula 3.2: Errore erlatiboa

espresioa duen. Demagun $C_1(n_1) = 6n_1 + 2n_1^2 + 2n_1^3$, $C_2(n_2) = 2n_2 + 2n_2^2 + 2n_2^3$, $C_3(n_3) = n_3 + n_3^2 + 3n_3^3$ eta $C_4(n_4) = 8n_4 + 2n_4^3$. Ohartu adibide honentzat $C'(n)\mu$ politikak (3.4.1) politika baino errendimendu kaxkarragoa erakusten duela, baina bi politikak soluzio optimotik oso gertu daude.

3.5 Errendimendu optimoa asintotikoki

Sekzio honetan Whittle indize politika asintotikoki optimoa den eremuak aztertuko dira kapitulo honetan aztergai den eredurako. 3.5.1. Sekzioan Whittle indizea optimoa den eztabaidatu da zerbitzari anitzdun sistemarako, eta 3.5.2. Sekzioan Whittle indize politika optimoa dela frogatuko da trafikoa arina eta trafikoa geldoa den kasuetan.

3.5.1 Zerbitzari anitzeko sistema

Mantentze kostuak linearrak direnean, zerbitzari-anitzeko sistemarako indize politika optimoa izatea zuzenean [92]-tik ondorioztatzen da. Demagun M zerbitzari daudela eta iritsiera tasa k klaseko bezeroentzat $M\lambda_k$ dela. Izan bedi W_k egoerarekiko askea den indizea (3.3.1) ekuazioan emanda bezela. [92, 6.2. Proposizioa]-ko emaitzak erakusten du Whittle indize politika (WI), non zerbitzari bakoitzak uneoro W_k ez negatiboa eta handiena duen bezeroa zerbitzatzen duen, asintotikoki optimoa da hurrengo zentzuan: edozein ϕ politikarako

$$\lim_{M\to\infty} \mathcal{C}^{WI}(M) \leq \liminf_{M\to\infty} \overline{\mathcal{C}}^{\phi}(M),$$

non $\mathcal{C}^{WI}(M)$ Whittle indizeak eragindako batezbesteko kostua den, eta $\overline{\mathcal{C}}^{\phi}(M)$ ϕ politikak eragindako batezbesteko kostua M zerbitzari daudenean sisteman.

Mantentze kostuak orokorrak direnean, ezin froga daiteke indize politika asintotikoki optimoa izatea. Halere, batek itxaron dezake hainbat baldintza betez gero hurrengoa betetzea. Demagun M zerbitzari daudela eta x_kM ilara non k klaseko bezeroak λ_k iritsiera tasa duten edozein $k = 1, \ldots, K$-tarako.[2] Ilara bakoitza gehienez zerbitzari batez zerbitzatu daiteke.

Bandit terminologiari dagokionez x_kM k klaseko *bandit* daudela esan daiteke non *bandit* bakoitzaren egoerak (bezero kopuruak) $E = \{0, 1, \ldots\}$ multzoan hartzen dituen balioak, eta planifikatzaile batek zein M *bandit* egingo diren aktibo erabaki behar du (eta zein M ilara zerbitzatu). E egoera espazioa finitua balitz, [92, 95]-ko emaitzak (hainbat baldintza betetzen direla ziurtatuz) Whittle indizea asintotikoki optimoa dela froga daiteke $M \to \infty$ den heinean. Halere, egoera espazioa infinitua den kasuan, eredu honen kasua den bezela, ez da emaitzarik ezagutzen.

[2] Hau k klaseko x_kM fluxuko sistema gisa ikus daiteke non bezero berriak iristen diren.

3.5.2 Trafiko arina eta trafiko geldoa

Trafiko arina eta trafiko geldoa, iritsiera tasa 0-runtz eta ∞-runtz doan kasuen adierazpide dira, hurrenez hurren. Limiteak hartu ahal izateko iritsiera tasak moldatuko dira klase bakoitzaren trafiko proportzioa finko mantentzeko. Hori egin ahal izateko, $\lambda_k = \gamma_k \lambda$ definitu, non λ iritsiera tasa totala den, eta $\sum_{k=1}^{K} \gamma_k = 1$.

Sekzio honetan 2.5. Sekzioan proposatu den metodoa erabiliko da Whittle indize politika asintotikoki optimoa dela frogatzeko trafikoa arina eta trafikoa geldoa diren kasuetan. Trafiko arineko eremuan, sistema hutsik aurkitzen da ia uneoro eta gehienez bezero bat aurkitzen da sisteman. Honek $\lambda \to 0$ den heinean, $REL(0)$ jatorrizko problemarentzat onargarria bilakatzen dela inplikatzen du, hau da, $REL(0) \in \mathcal{U}$. Trafikoa geldoa denean, W-ren aukeraketa ona eginez, Whittle indizearen politikapean, (2.3.1) baldintza betetzen da berdintasunez, eta $REL(W) \in \mathcal{U}$.

3.8. Proposizioan trafiko arinean indize politika asintotikoki optimoa izatea frogatu da eta 3.9. Proposizioan trafiko geldoari dagokion emaitza aurki daiteke. Bi frogapenak 3.7.7. eta 3.7.8. Eranskinetan aurki daitezke, hurrenez hurren.

3.8 Proposizioa. *Demagun $C_k(0,0) \geq C_k(0,1)$, edozein k-tarako. Whittle indize politika (WI) asintotikoki optimoa da trafiko arinean, hau da,*

$$\lim_{\lambda \downarrow 0} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} = 0,$$

with $\lambda_k = \lambda\gamma_k$, $\sum_{k=1}^{K} \gamma_k = 1$.

3.9 Proposizioa. *Demagun mantentze kostuak linealak direla eta beraz $W_k(n) = w_k$ Whittle indizea konstantea dela. Demagun existitzen dela $\bar{k} \in \{1, \ldots, K\}$ non*

$$W_k < W_{\bar{k}},$$

edozein $k \neq \bar{k}$-rako. Orduan (WI) Whittle indize politika asintotikoki optimoa da trafikoa geldoa denean, hau da,

$$\lim_{\lambda \to \infty} \frac{\mathcal{C}^{WI} - \mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} = 0,$$

non $\lambda_k = \lambda\gamma_k$, $\sum_{k=1}^{K} \gamma_k = 1$.

3.9. Proposizioaren frogapenetik ikus daiteke, $\bar{k}$ klaseari lehentasuna ematen dion edozein politika dela optimoa $\lambda \uparrow \infty$ den heinean.

3.6 Emaitza numerikoak

Sekzio honetako helburua Whittle indizeak zein eremuetan duen errendimendu ona erakustea da (non indizea (2.3.6) ekuazioak definitzen duen). Aztergai diren kostu funtzioak $C_k(n_k, a) = C_k(n_k)$ edo $C_k(n_k, a) = C_k((n_k - a)^+)$ dira, hau da, mantentze kostua k klaseko ilaran daude edo sisteman dauden bezero kopuruaren funtzio bat denean, hurrenez hurren. Beraz, $\tilde{C}_k(n_k, a)$ funtzioa, $C_k(n_k) + \delta_k\theta_k n_k$ edo $C_k((n_k - a)^+) + \delta_k\theta_k(n_k - a)^+ + \delta'_k\theta'_k \min(a, n_k)$ bezela idatz daiteke, hurrenez hurren.

3.6.1. Sekzioan Whittle indize politikaren egitura, politika optimoaren egiturarekin erkatu da, numerikoki. 3.6.2. Sekzioan indize politiken errendimendua numerikoki erkatu da soluzio optimoarekin.

3.6.1 Politika ezberdinen egitura

Indize politika ezberdinen egitura soluzio optimoaren egiturarekin erkatu da mantentze kostu linealak edo ganbilak direnean.

Mantentze kostu linearrak

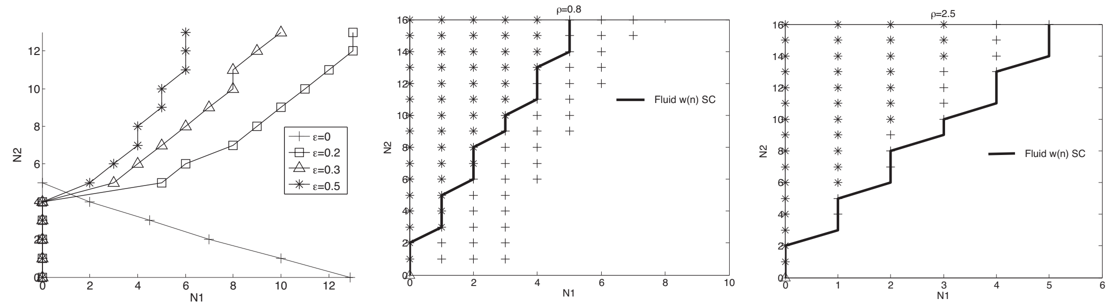
Value iteration algoritmoa erabiliz, ikusi 1.3.3. Sekzioa, parametro ezberdin askotarako, politika optimoa kalkulatu da, mantentze kostu linealak diren kasurako, eta hurrengo egitura jarraitzen du: $(N_1, \ldots, N_K)$ jatorriarekiko gertu dagoenean (non N_i i klaseko bezero kopuruaren adierazle den), optimoa da klaseen lehentasuna $\tilde{c}\mu$ politikaren araberakoa izatea, eta jatorritik gertu ez dagoenean $\tilde{c}\mu/\theta$ politikaren araberakoa izatea, non $\tilde{c}_k := c_k + \delta_k\theta_k$, ikusi 3.1. Irudia (ezkerraldean) non $\epsilon = 0$, hurrengo sekzioan azalduko den bezela. Beraz, Whittle indizeak ($\tilde{c}\mu/\theta$ balioa hartzen duena kostuak linearrak direnean) akzio optimoa atzematen du egoera jatorritik oso gertu ez dagoenean.

Mantentze kostu orokorrak

Indize politiken egitura eztabaidatzeko mantentze kostu orokorretarako, bi bezero klaseko sistemetan oinarrituko da analisia ($K = 2$). (N_1, N_2) egoeran, Whittle indize politikak hartzen duen akzioa $W_k(N_k)$ handiena duen klasea zerbitzatzea da. $W_k(N_k)$ funtzio ez-beherakorra denez, trukatze-funtzio (SC) gorakor bat existitzen dela inplikatzen du, non (N_1, N_2) SC-ren azpitik dagoenean, Whittle indize politikak 1 klasea zerbitzatzea erabakitzen duen eta (N_1, N_2) SC-ren gainetik badago 2 klasea zerbitzatzea. Ohartu kostuak linealak direnean trukatze-funtzio hau ardatz horizontal edo bertikalean kolapsatzen dela.

Value iteration algoritmoa erabiliz soluzio optimoa trukatze-funtzio bat dela ikusi da. Adibidez, 3.1. Irudian (ezkerraldean) soluzio optimoaren trukatze-funtzioa irudikatu da kostuak $C_1(n) = n + \epsilon n^2$ eta $C_2(n) = n$ diren kasuan ($\theta = \theta'$ eta $\lambda = [9, 10], \mu = [14, 16], \theta = [2, 0.05], \delta = [4, 0.3]$ parametroentzat). $\epsilon = 0$ denean, trukatze-funtzioa beherakorra da, zeinak kostu linealetarako 3.6.1. Sekzioan azaldu den politika optimoaren errendimendua deskribatzen duen. ϵ positiboa bilakatzen den heinean, trukatze-funtzioa gorakorra bilakatzen da. Ez hori bakarrik, ϵ handitzen den heinean, eta beraz 1 klaseko bezeroen kostu kuadratikoa handitzen da, orduan, 1 klaseak lehentasuna eskuratzen du eremu handiago batean.

Orain Whittle indize politikak eta politika optimoak hartzen dituzten akzioak erkatuko dira. Kostu kuadratikoak dituen adibide bat izango da aztergai, non $C_1(n) = (c_{11} + \delta_1\theta_1)n + c_{21}n^2$ eta $C_2(n) = (c_{12} + \delta_2\theta_2)n + c_{22}n^2$, eta parametroak $\theta = \theta'$, $\mu = [15, 18]; \theta = [4, 7]; c_1 = [1, 4]; c_2 = [2, 1]$ eta $\delta = [8, 6.5]$ diren. 3.1. Irudian (erdian eta eskuinaldean) akzio optimoak irudikatu dira (*value iteration* bidez lortuak) lan kargak $\rho = 0.8$ eta $\rho = 2.5$ direnean, hurrenez hurren, eta Whittle indize politiken akzioakin erkatu dira. Politika optimoa trukaze-funtzio batez adieraz daiteke. Gainera, politika optimoa $W(n)$ Whittle indize politikaren baliokidea bilakatzen da lan karga handitzen den heinean. Indize fluido politika, $w(n)$ ere irudikatu da eta ikusi errendimendua oso ona dela. Indize fluidoa $w(n)$, 4. Kapituluan aurkeztuko da.



Irudia 3.1: Ezkerraldea: Trukatze-funtzioa politika optimoarentzat mantentze kostua aldatzen den heinean (linealetik kuadratikora). Erdian eta eskuinaldean: Politika optimoaren akzioak, $W(n)$ indize politikarenak eta indize fluido politikaren mantentze kostuak kuadratikoak diren kasurako. "+" ikurdun eremua: $W(n)$ politikak 1 klasea zerbitzatzen du 2 klasea zerbitzatzea delarik akzio optimoa. "*" ikurdun eremua: $W(n)$ politikak 2 klasea zerbitzatzen du, akzio optimoa dena, eta hutsik utzitako eremuan $W(n)$ 1 klasea zerbitzatzen du, akzio optimoa horixe delarik.

3.6.2 Errendimenduaren ebaluazioa

Sekzio honetan indize politiken errendimendua balioztatuko da. Hau egin ahal izateko soluzio optimoaren eta indize politikak eragindako batez besteko kostuaren arteko errore erlatiboa kalkulatuko da. Horretarako, *value iteration* algoritmoa erabili da.

3.4. Sekzioan ikusi da (3.4.1) indizetzat duen indize politikak errendimendu ona erakusten duela M/M/1 ilara klase-anitzetan (bezeroen uzterik gertatzen ez denean). Kapitulu horretan kostu kubikoak eta 4 bezero klase kontsideratu dira eta $Gc\mu$ indize politika (3.4.1) indize politikarekin erkatu da, bertan ikusi ahal izan da azken politika honek $Gc\mu$ politikak baino errendimendu hobea duela.

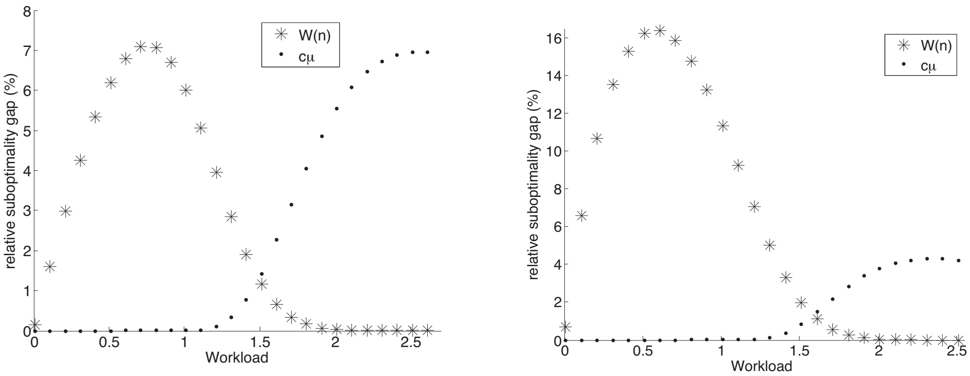
Sekzio honetan bezeroen uzteak dituzten sistemak izango dira aztergai. Hurrengo indizeen ebaluazioa egingo da: (i) $W(n)$ Whittle indize politika ((2.3.6) ekuazioak emana), (ii) $W^{\infty}(n)$ egoera handietarako Whittle indizea, 3.4. Proposizioan proposatu dena, eta (iii) $w(n)$ indize fluidoa, 4. Kapituluan aztertuko dena. Indize politika hauek uzterik gabeko ilara klase anitzentzat lortu diren bi indize politikekin erkatuko dira: $Gc\mu$ politika eta (3.4.1) ekuazioari dagokion indize politika, zeina $W(n)$-ren hurbilketa bat den θ 0-runtz doan kasurako. Bi egoera ezberdin aztertuko dira: (1) ρ lan karga aldatzen den kasua, eta (2) θ aldatzen den kasua.

Lan karga aldatuz

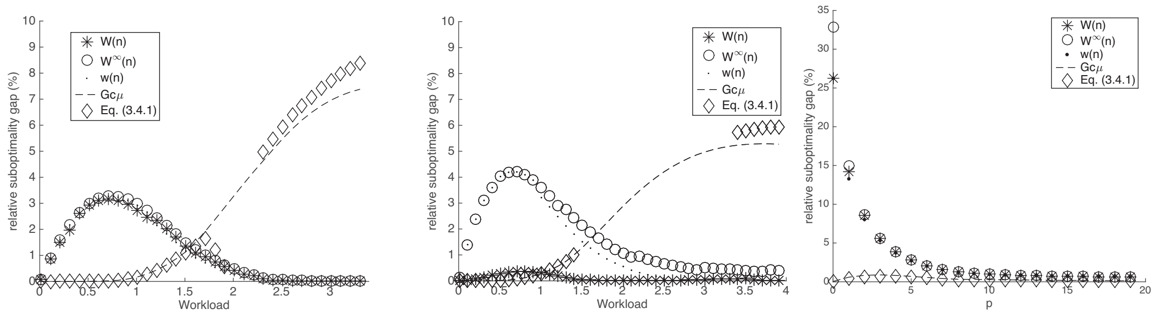
Sekzio honetako helburua indize politiken jokaera aztertzea da lan karga ezberdinetarako.

Mantentze-kostu linealen adibidea ($\theta = \theta'$) Finkatu $C_k(n, a) = c_k n$, $\mu = [15, 25]$, $\theta' = \theta = [4, 2]$, $c = [1, 1]$, $\delta = [5, 3.2]$ eta demagun $\rho = \sum_{k=1}^{2} \lambda_k/\mu_k$-k $[0, 2.6]$ tartean hartzen dituela balioak, non $\lambda_1/\mu_1 = \lambda_2/\mu_2$. Mantentze-kostu linealetarako, $W(n), W^{\infty}(n)$ eta $w(n)$ indize politikak $\tilde{c}\mu/\theta$ indize politikara sinplifikatzen dira eta $Gc\mu$ eta (3.4.1) $\tilde{c}\mu$ indize politikara, non $\tilde{c}_k = c_k + \delta_k\theta_k$.

Mantentze-kostu linealen adibidea ($\theta \neq \theta'$)**:** Finkatu $C_k(n, a) = c_k(n-a)^+$, $\mu = [15, 25]$, $\theta = [4, 2]$, $\theta' = [3, 2]$, $c' = c = [1, 1]$, $\delta = [5, 3.2]$, $\delta' = [2, 1]$ eta demagun $\rho = \sum_{k=1}^{2} \lambda_k/\mu_k$-k $[0, 2.6]$ tartean hartzen dituela balioak, non $2\lambda_1/\mu_1 = \lambda_2/\mu_2$. Mantentze-kostu linealetarako eta $\theta \neq \theta'$, $W(n), W^{\infty}(n)$ eta $w(n)$ indize politikak $\tilde{c}(\mu+\theta')/\theta - \tilde{c}'$ indize politikara sinplifikatzen dira, eta $Gc\mu$ eta (3.4.1) $\tilde{c}\mu$ indize politikara, non $\tilde{c}_k = c_k + \delta_k\theta_k$.



Irudia 3.2: Ezkerraldean: errore erlatiboa mantentze kostu-linearrentzat, ρ handitzen den heinean eta $\theta = \theta'$. Eskuinaldean: errore erlatiboa mantente-kostu linearrentzat ρ handitzen den heinean eta $\theta \neq \theta'$.



Irudia 3.3: Ezkerraldean: errore erlatiboa mantentze-kostu linealentzat, ρ handitzen den heinean. Erdialdean: errore erlatiboa mantentze-kostu kuadratikoentzat, ρ handitzen den heinean. Eskuinaldean: errore erlatiboa mantentze-kostu kuadratikoentzat, p ($\theta_i = p\epsilon_i, i \in \{1,2\}$) aldatzen den heinean.

3.2. Irudian bi kasuetan ikusi da $\tilde{c}\mu$ optimoa dela lan karga txikia denean, eta $W(n)$ indizearen errendimendua ia optimoa da lan karga handia denean, 3.9. Proposizioan aipatu bezala. 3.6.1. Sekzioan eztabaidatu den legez, jatorritik urruti dauden egoeretan, akzio optimoa $\tilde{c}\mu/\theta$ indizearen arabera zerbitzatzea da, hori baita batez bestean prozesua biziko den eremua. Horrek azaltzen du $\tilde{c}\mu/\theta$ eta $\tilde{c}(\mu+\theta')/\theta - \tilde{c}'$ indize politikek errendimendu ona erakustea kasu honetan. Lan karga txikia denean, bezeroen uzteen eragina ez da hain garrantzitsua eta beraz $\tilde{c}\mu$ indizeak errendimendu ona erakusten du.

Mantentze-kostu kuadratikoen adibidea ($\theta = \theta'$) Demagun hurrengo parametroak kontsideratu direla: $\mu = [15, 18]$, $\theta' = \theta = [4, 7]$, $c_1 = [1, 4]$, $c_2 = [2, 1]$, $\delta = [8, 6.5]$ eta demagun λ aldatuz doala, $\lambda_1/\mu_1 = \lambda_2/\mu_2$ bermatuz. Demagun kostu kuadratikoak $C_1(n) = (c_{11} + \delta_1\theta_1)n + c_{21}n^2$ eta $C_2(n) = (c_{12} + \delta_2\theta_2)n + c_{22}n^2$ direla. Ikusi 3.3. Irudian errore erlatiboa eta 3.3. Taulan errore absolutua.

Ikusi ahal izan da lan karga txikia denean $Gc\mu$ eta (3.4.1) indize politikak errendimendu ona dutela. Halere, lan karga handitzen den heinean, θ-rekiko askeak diren indize politika hauen errore erlatiboa handitzen da, $W(n)$ Whittle indize politika, $W^\infty(n)$ Whittle indize politika egoera handientzat eta $w(n)$ indize fluido politika ia optimo bilakatzen dira. 3.3. Taulan ikusi daiteke optimorainoko konbergentzia oso azkarra dela $W(n), W^\infty(n)$ eta $w(n)$ indize politiken errore absolutua ($\mathcal{C}^{WI} - \mathcal{C}^{OPT}$) 10^{-4} mailakoa dela $\rho = 5.25$ den heinean. Bestalde, bai (3.4.1) eta $Gc\mu$ indize politikek oso errendimendu kaxkarra erakusten

Lan-karga	1	1.5	2	2.5	3	3.5	5.25
$W(n)$	1.3089	1.4608	0.8055	0.1094	0.0185	0.0065	0.00017
$W^{\infty}(n)$	1.4028	1.5596	0.8902	0.1732	0.0614	0.0329	0.0007
$w(n)$	1.3823	1.2885	0.5534	0.0026	0.0771	0.0904	0.0004
(3.4.1)	0.0409	0.7327	0.8010	11.2134	20.5851	28.3926	50.0996
$Gc\mu$	0.0409	0.7483	3.9951	10.4111	18.7237	25.0454	42.5645

Taula 3.3: $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ errore absolutua, 3.3. Irudiko (ezkerraldean) adibideari dagokiona.

Lan-karga	1	1.5	2.5	3	3.5	5.25	7.25	10	16
$W(n)$	0.1332	0.0664	0.0098	0.1260	0.2874	0.2448	0.1404	0.0486	0.0061
$W^{\infty}(n)$	1.4817	1.9167	1.4429	1.1485	1.4243	1.7296	1.4784	0.7977	0.1012
$w(n)$	1.4817	1.4157	0.3397	0.0382	0.1288	0.5125	04383	0.1542	0.0093
(3.4.1)	0.0720	-	-	-	19.3226	35.5180	48.5766	66.1024	91.4859
$Gc\mu$	0.0720	0.7896	7.7697	12.8528	17.6942	31.1417	43.3748	59.7161	99.4344

Taula 3.4: $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ errore absolutua 3.3. Irudiari dagokion adibiderako (erdialdean).

dute lan karga handietarako. Beraz, tesi honetan garatu diren indize politikak lan karga handietarako dira egokiak, eta hau da praktikan interesgarria den kasua.

$\rho = 2$ inguruan (3.4.1) indize politikak erakusten duen jauzia $\lambda_k = \mu_k$ inguruan balioak definitu gabe egotearen ondorio da.

Mantentze-kostu kuadratikoen adibidea ($\theta \neq \theta'$)**:** Demagun hurrengo parametroak kontsideratu direla: $\mu = [15, 18]$, $\theta = [4, 7]$, $\theta' = [3, 4]$, $c_1 = [1, 4]$, $c_2 = [2, 1]$, $\delta = [8, 6.5]$, $\delta' = [7, 7]$ eta λ aldatzen doala $2\lambda_1/\mu_1 = \lambda_2/\mu_2$ bermatuz. Demagun kostuak kuadratikoak direla $\tilde{C}_1(n, a) = c_{11}(n-a)^+ + c_{21}((n-a)^+)^2 + \delta_1\theta_1(n-a)^+ + \delta_1'\theta_1' a$ eta $\tilde{C}_2(n, a) = c_{12}(n-a)^+ + c_{22}((n-a)^+)^2 + \delta_2\theta_2(n-a)^+ + \delta_2'\theta_2 a$. Ikusi 3.2. Irudia errore erlatiborako eta 3.4. Taula errore absoluturako.

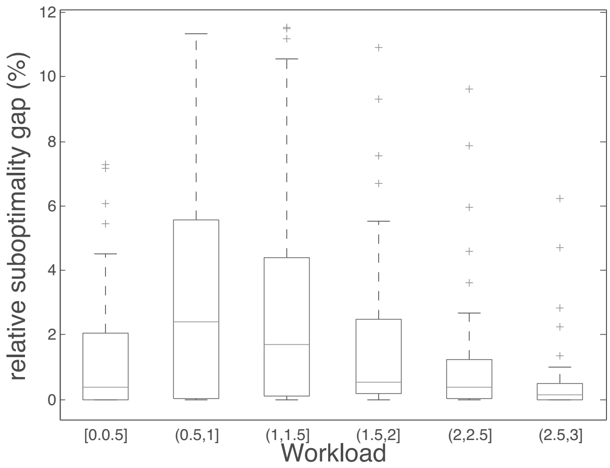
Ikusi da lan karga txikia denean $Gc\mu$ eta (3.4.1) indize politikek errendimendu ona erakusten dutela. Adibide honetan, Whittle indize politikak ere ia errendimendu optimoa erakusten du lan karga txikierako, ordez $W^{\infty}(n)$ eta $w(n)$ indize politikek errendimendu kaxkarragoa dute. Lan karga handitzen den heinean, $W(n)$ Whittle indize politika eta $w(n)$ indize fluido politika ia optimoak bilakatzen dira. Halere, adibide honetan optimorako konbergentzia errore absolutuari dagokionez askoz mantsoagoa da aurreko adibideetan baino. $\mathcal{C}^{WI} - \mathcal{C}^{OPT}$ errore absolutua 10^{-3} mailakoa da $W(n)$ eta $w(n)$ indize politikentzat eta 10^{-1} mailakoa $W^{\infty}(n)$ indize politikarentzat $\rho = 16$ denean. Fenomeno hau ulertzeko ohartu politika optimoan batez bestean bizi den eremuan 2 klasea zerbitzatzea erabakitzen du, baina indize politikek 1 klasea zerbitzatzea erabakitzen dute. Lan karga handitzen den heinean fenomeno hau desagertu egiten da.

$\rho = [1.5, 3]$ tartearen (3.4.1) indize politikak erakusten duen jauzia $\lambda_k = \mu_k$ inguruan balioak definitu gabe egotearen ondorio da.

Uzte tasak aldatuz

Sekzio honetan indize politiken errendimendua aztertuko da uzte tasak aldatzen diren heinean.

Mantentze-kostu linealak: Kasu honetan, goian aipatu diren bost indize politikak $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu$ indize politiketara sinplifikatzen dira, 3.6.2. Sekzioan azaldu den legez. $\theta_k \to 0$-runtz doan heinean, emaitza numerikoek erakutsi dute $\tilde{c}\mu$ indize politikak errendimendu optimoa duela, eta $\tilde{c}\mu/\theta$ indize politikak errendimendu oso kaxkarra izan dezakeela uzte tasak oso txikiak direnean. Jakina da, uzterik ez duten



Irudia 3.4: Whittle indize politikaren errore erlatiboa, zoriz sortu diren parametroentzat. Kaxen ertzek 25. eta 75. pertzentilak erakusten dituzte, kaxen barruko marrak batez besteko balioa erakusten dute eta "+-k balio ez-ohikoak erakusten ditu.

sistemetan $\tilde{c}\mu$ politika dela optimoa lan karga txiketarako (M/M/1 ilarentzat $c\mu$ erregela ospetsuaren baliokide dena). Halere, $\tilde{c}\mu/\theta = (c + \delta\theta)\mu/\theta$ indize politikak kontrako akzioa har dezake θ-ren balioak oso txikiak direnean, $\tilde{c}\mu/\theta$ indizeak errendimendu kaxkarra izatea azaltzen du honek θ-k balio txikiak dituenean.

Mantentze-kostu kuadratikoak: Demagun bi bezero klase daudela sisteman eta mantentze kostuak kuadratikoak direla $C_1(n) = \tilde{c}_{11}n + c_{21}n^2$ non $\tilde{c}_{11} = (c_{11} + \delta_1\theta_1)$, eta $C_2(n) = \tilde{c}_{21}n + c_{22}n^2$, non $\tilde{c}_{21} = (c_{21} + \delta_2\theta_2)$ eta hurrengo parametroak dituela sistemak: $\lambda = [4, 5]$, $\mu = [15, 17]$, $c_1 = [1, 4]$, $c_2 = [5, 1]$, $\delta = [2, 3]$, $\theta_1 = \epsilon_1 p$ eta $\theta_2 = \epsilon_2 p$, non $\epsilon_1 = 0.05$, $\epsilon_2 = 0.01$ eta p aldatuz doala. Beraz, $\rho = \sum_k \rho_k < 1$ sistema egonkorra dela bermatzeko $\theta_k \to 0$ den heinean.

3.3. Irudian (eskuinaldean) errore erlatiboa irudikatu da p 0-tik 200-era aldatzen den heinean, beraz θ_1 eta θ_2 $[0, 10]$ eta $[0, 2]$ tarteetan aldatzen dira, hurrenez hurren. Ikusi θ-rekiko askeak diren indize politikak %25-eko errore erlatiboa erakusen dutela $p = 0$ inguruan. θ handitzen den heinean, errore hori azkar desagertzen da. θ-rekiko askeak diren indize politikek errendimendu ona dute karga txikiko kasuak aztertu direlako hemen.

Zorizko laginen adibidea

Sekzio honetan 360 lagin ezberdin hartu dira kontutan, zeinetarako $\lambda_k, \mu_k, \theta'_k, \theta_k$ eta $c_k = [c_{k1}, c_{k2}]$-ren balioak zorizkoak diren edozein $k = 1, 2$-rako. c_k aurreko adibidean definitu da. Whittle indize politikaren errore erlatiboa kalkulatu da, ikusi 3.4. Irudia. 3.4. Irudia kaxa-diagrama bat da, non kaxa bakoitzak 60 laginen errore erlatiboa gordetzen duen. 0.5-eko tamaina duten tarteetan batu dira emaitzak. Ohartu Whittle indize politikaren errendimendua ia optimoa dela lan karga handietarako, eta errendimendu kaxkarrena $(0.5, 1]$ tarteko balioek erakusten dute.

3.7 Eranskina

3.7.1 3.1. Proposizioaren frogapena

3.1. Proposizioaren helburua $\phi = n$ atari-politika optimoa dela (2.3.3) problemarentzat frogatzea da. Hori egin ahal izateko, V balio funtzioa ganbila dela ikustea falta da. Horretarako $V^L(m)$ sistema moztuaren balio funtzioa ($L > 1$ parametroak moztua) ganbila dela ikusiko da. Hori egin ostean [25, 3.1. Teorema] emaitza dela eta $V^L(m) \to V(m)$ froga daiteke $L \to \infty$ den heinean eta beraz, V^L ganbila izateak edozein L-tarako V ganbila izatea inplikatzen du. [25, 3.1. Teorema] aplikatu ahal izateko hainbat baldintza betetzen direla ikusi behar da. Beraz, lehenik eta behin baldintza horiek betetzen direla ikusiko da eta gero V^L ganbila dela frogatuko da.

[25, 3.1. Teorema] betetzeko behar diren baldintzak

Lehenik eta behin hurrengo definizioa aurkeztuko da:

3.1 Definizioa. *$f : E \longrightarrow \mathbb{R}_+$ funtzioa momentu-funtzioa dela esango da $E_r \uparrow E$ existitzen bada, $r \to \infty$-tarako, non $\inf\{f(m) : m \notin E_r\} \to \infty$ $r \to \infty$ den heinean. (E egoera espazioa da).*

Definitu $q^{\phi,L}(m, m-1) = \mu S^\phi(m) + \theta' S^\phi(m) - \theta(m - S^\phi(m))$, eta gogoratu $q^{\phi,L}(m, m+1) = \lambda\left(1 - \frac{m}{L}\right)$. [25, 3.1. Teorema] betetzeko behar diren baldintzak:

1. $f : \mathbb{N} \cup \{0\} \longrightarrow \mathbb{R}_+$ momentu funtzio bat existizen da, eta $\alpha, \beta > 0$ eta $\tilde{M} > 0$ konstanteak zeinetarako

$$\sum_{\tilde{m}=0}^{\infty} q^{\phi,L}(m, \tilde{m}) f(\tilde{m}) \leq -\alpha f(m) + \beta \mathbf{1}_{\{m<\tilde{M}\}}(m), \text{ edozein } \phi, L,$$

   non ϕ jarraitu den politika den, L mozte parametroa, eta $q^{\phi,L}(m, \tilde{m})$ trantsizio tasa m-tik $\tilde{m}$-ra ϕ politikapean eta L parametroarekiko.

2. $(S^\phi(m), L) \mapsto q^{\phi,L}(m, \tilde{m})$ eta $(S^\phi(m), L) \mapsto \sum_{\tilde{m}} q^{\phi,L}(m, \tilde{m}) f(\tilde{m})$ funtzio jarraiak dira $S^\phi(m)$ eta L-n edozein m eta $\tilde{m}$-rentzat.

Definitu $f(m) := \mathrm{e}^{\epsilon m}$, non $\epsilon > 0$. Eraiki $E_r = \{0, \ldots, r\}$ non E_r finitua den, eta $E_r \uparrow \mathbb{N} \cup \{0\}$ $r \to \infty$ doan heinean eta $\inf\{f(m) : m \notin E_r\} \to \infty$. Helburua da $\epsilon, \alpha, \tilde{M} > 0$-rik existitzen diren ikustea, zeinentzat

$$\sum_{\tilde{m}=0}^{\infty} q^{\phi,L}(m, \tilde{m}) f(\tilde{m}) \leq -\alpha f(m), \text{ edozein } m \geq \tilde{M}\text{-tarako},$$

hau da,

$$\begin{aligned}
&\lambda\left(1 - \frac{m}{L}\right) \mathrm{e}^{\epsilon(m+1)} + \left((\mu + \theta') S^\phi(m) + \theta(m - S^\phi(m))\right) \mathrm{e}^{\epsilon(m-1)} \\
&- \left(\left(1 - \frac{m}{L}\right) + (\mu + \theta') S^\phi(m) + \theta(m - S^\phi(m))\right) \mathrm{e}^{\epsilon m} \leq -\alpha \mathrm{e}^{\epsilon m}, \text{ edozein } m \geq \tilde{M}\text{-tarako.}
\end{aligned}$$

Kalkulu batzuen ostean

$$\lambda\left(1 - \frac{m}{L}\right)(\mathrm{e}^{\epsilon} - 1) + \left((\mu + \theta' - \theta) S^\phi(m) + \theta m\right)(\mathrm{e}^{-\epsilon} - 1) \leq -\alpha, \text{ edozein } m \geq \tilde{M}\text{-tarako,}$$

lortzen da. Ohartu $\lambda(1-m/L)(\mathrm{e}^{-\epsilon}-1)$ goitik bornatu daitekeela konstante batez, κ_1, eta $(\mu+\theta'-\theta)S^{\phi}(m)(\mathrm{e}^{\epsilon}-1)$ goitik bornatu daiteke κ_2 konstantearekin. Bestalde, $\theta m(\mathrm{e}^{-\epsilon}-1)<0$. Beraz, $\tilde{M}$ aurki daiteke nahiko handia $-\theta m(\mathrm{e}^{-\epsilon}-1)\geq\kappa_1+\kappa_2$ bete dadin edozein $m\geq\tilde{M}$-tarako. Honek (1) baldintza bermatzen du.

(2) baldintza, *i.e.*, $(S^{\phi}(m),L)\mapsto q^{\phi,L}(m,\tilde{m})$ eta $(S^{\phi}(m),L)\mapsto\sum_{\tilde{m}}q^{\phi,L}(m,\tilde{m})f(\tilde{m})$ funtzioak jarraiak izatea L eta $S^{\phi}(m)$ trantsizio tasen definizioa dela-eta betetzen da.

V^L-ren ganbiltasuna

Notazioa sinplifikatzearren $\omega:=\mu+\theta'-\theta$ definituko da frogapenean zehar. Orokortasunik galdu gabe demagun $\lambda+\mu+\theta'+\theta L=1$. Edozein $n\in\{0,1,\ldots,L\}$-tarako definitu $V_t^L(n)$ non $V_0^L(n)=0$ eta

$$\begin{aligned}V_{t+1}^L(n)=&\lambda\left(1-\frac{n}{L}\right)V_t^L(\min\{n+1,L\})\\&+\min\{-W+\tilde{C}(n,0)+\omega V_t^L(n),\tilde{C}(n,1)+\omega V_t^L((n-1)^+)\}\\&+\theta nV_t^L((n-1)^+)+\lambda\frac{n}{L}V_t^L(n)+(L-n+1)\theta V_t^L(n).\end{aligned}$$

V_t^L funtzio ganbila dela frogatuko da edozein $n\leq L-1$-tarako, hau da,

$$2V_t^L(n)\leq V_t^L((n-1)^+)+V_t^L(n+1),\ \text{non } n\leq L-1. \tag{3.7.1}$$

V_t^L funtzio ganbila izateak, edozein t-rako, V^L ganbila izatea inplikatzen du eta horrekin frogapena amaitzen da.

V_t^L ganbila dela frogatzeko $V_t^L(\cdot)$ funtzio ez-beherakorra dela frogatu behar da. Frogapena indukzioz egin daiteke: $V_0^L(n)=0$ ez-beherakorra da $t=0$-rentzat, onartu $V_t^L(n)$ ez-beherakorra dela eta frogatu

$$V_{t+1}^L(n+1)-V_{t+1}^L(n)\geq 0 \text{ edozein } n\leq L-1\text{-tarako.} \tag{3.7.2}$$

Lehenik eta behin $V_{t+1}^L(n+1)-V_{t+1}^L(n)$-n λ-z biderkatutako gaiak izango dira aztergai, hau da,

$$\begin{aligned}&\lambda\left(1-\frac{n+1}{L}\right)V_t^L(\min\{n+2,L\})+\lambda\frac{n+1}{L}V_t^L(\min\{n+1,L\})\\&-\lambda\left(1-\frac{n}{L}\right)V_t^L(\min\{n+1,L\})-\lambda\frac{n}{L}V_t^L(n)\\&\geq\lambda\left(1-\frac{n+1}{L}\right)(V_t^L(\min\{n+2,L\})-V_t^L(\min\{n+1,L\}))\\&\quad+\lambda\frac{n}{L}(V_t^L(\min\{n+1,L\}-V_t^L(n)))\geq 0,\end{aligned}$$

non azken inekuazioa $V_t^L(n)$ ez-beherakorra izateak inplikatzen duen. Orain $V_{t+1}^L(n+1)-V_{t+1}^L(n)$-n θ-z bidertakutako gaiak dira aztergai, hau da,

$$\begin{aligned}&\theta(n+1)V_t^L(n)+(L-n-1)\theta V_t^L(\min\{n+1,L\})-\theta nV_t^L((n-1)^+)-(L-n)\theta V_t^L(n)\\&\geq\ \theta n(V_t^L(n)-V_t^L((n-1)^+))+(L-n-1)(V_t^L(\min\{n+1,L\})-V_t^L(n))\geq 0,\end{aligned}$$

non, berriz ere, azken inekuazioa $V_t^L(n)$ ez-beherakorra izateak inplikatzen duen edozein $n \leq L-1$. Azkenik, $V_{t+1}^L(n+1) - V_{t+1}^L(n)$-ko min gaiak dira aztergai. Tribiala da

$$\begin{aligned}
&\min\{-W + \tilde{C}(\min\{n+1, L\}, 0) + (\mu+\theta')V_t^L(\min\{n+1, L\}),\\
&\qquad \tilde{C}(\min\{n+1, L\}, 1) + (\mu+\theta')V_t^L(n)\}\\
&-\min\{-W + \tilde{C}(n, 0) + (\mu+\theta')V_t^L(n),\\
&\qquad \tilde{C}(n, 1) + (\mu+\theta')V_t^L((n-1)^+)\} \geq 0,
\end{aligned}$$

dela. Azken hau $\tilde{C}$ eta V_t^L ez-beherakorra izateak inplikatzen dute.

Honek (3.7.2) frogatzen du eta beraz, $V_t^L(n)$ ez-beherakorra dela frogatu da.

(3.7.1). Ekuazioa $n=0$ denean $V_t^L(\cdot)$ ez-beherakorra izateak inplikatzen du. Frogapenaren gainontzekoan (3.7.1). Ekuazioa $n \geq 1$-erako frogatuko da.

(3.7.1) ganbiltasuna t-ren gaineko indukzioa egiten frogatuko da. $V_0^L(n) = 0$ denez, $t=0$ denerako betetzen da. Orain onartu V_t^L ganbila dela. Edozein $1 \leq n \leq L-1$-tarako

$$\begin{aligned}
2V_{t+1}^L(n) = {}& 2\lambda\left(1-\frac{n}{L}\right)V_t^L(n+1) + 2\lambda\frac{n}{L}V_t^L(n) + 2\theta n V_t^L(n-1) + 2(L-n+1)\theta V_t^L(n)\\
&+ 2\min\{-W + \tilde{C}(n,0) + \omega V_t^L(n), \tilde{C}(n,1) + \omega V_t^L(n-1)\},
\end{aligned} \tag{3.7.3}$$

da. Azken hau $V_{t+1}^L(n-1) + V_{t+1}^L(n+1)$ baino txikiagoa edo berdina dela ikusi nahi da, zeina

$$\begin{aligned}
&\lambda\left(1-\frac{n-1}{L}\right)V_t^L(n) + \lambda\left(1-\frac{n+1}{L}\right)V_t^L(n+2) + \lambda\frac{n-1}{L}V_t^L(n-1) + \lambda\frac{n+1}{L}V_t^L(n+1)\\
&+ \theta(n-1)V_t^L((n-2)^+) + \theta(n+1)V_t^L(n) + (L-n+2)\theta V_t^L(n-1) + (L-n)\theta V_t^L(n+1)\\
&+ \min\{-W + \tilde{C}(n-1,0) + \omega V_t^L(n-1), \tilde{C}(n-1,1) + \omega V_t^L((n-2)^+)\}\\
&+ \min\{-W + \tilde{C}(n+1,0) + \omega V_t^L(n+1), \tilde{C}(n+1,1) + \omega V_t^L(n)\},
\end{aligned} \tag{3.7.4}$$

ekuazioak definitzen duen. Lehenik ea behin (3.7.3). Ekuazioan λ-z biderkatutako gaiak aztertuko dira eta

$$\lambda\left(1-\frac{n-1}{L}\right)V_t^L(n) + \lambda\left(1-\frac{n+1}{L}\right)V_t^L(n+2) + \lambda\frac{n-1}{L}V_t^L(n-1) + \lambda\frac{n+1}{L}V_t^L(n+1), \tag{3.7.5}$$

baino txikiagoa edo berdina dela frogatuko da. $1 \leq n < L-1$ bada, (3.7.3). Ekuazioan λ-z biderkatutako gaiak

$$\begin{aligned}
&2\left(1-\frac{n}{L}\right)V_t^L(n+1) + 2\frac{n}{L}V_t^L(n) = 2\left(1-\frac{n+1}{L}\right)V_t^L(n+1) + 2\frac{n}{L}V_t^L(n) + \frac{2}{L}V_t^L(n+1)\\
&\leq \left(1-\frac{n-1}{L}\right)V_t^L(n) - \frac{2}{L}V_t^L(n) + \left(1-\frac{n+1}{L}\right)V_t^L(n+2) + 2\frac{n}{L}V_t^L(n) + \frac{2}{L}V_t^L(n+1),
\end{aligned} \tag{3.7.6}$$

idatz daitezke. Azken inekuazioa V_t^L ganbila izateak inplikatzen du. Ganbiltasunak $2\frac{n-1}{L}V_t^L(n) \leq \frac{n-1}{L}(V_t^L(n-1) + V_t^L(n+1))$ inplikatzen du, eta beraz, (3.7.6) (3.7.5) baino txikiago edo berdina da. $n = L-1$ bada, $2(1-2/L)V_t^L(L-1) \leq (1-2/L)(V_t^L(L-2) + V_t^L(L))$ ziurtatzea nahikoa da, azken hau V_t^L ganbila izateak inplikatzen du.

θ-rekin biderkatutako gaietarako nahikoa da

$$\begin{aligned}&2nV_t^L(n-1)+2V_t^L(n)+2(L-n)V_t^L(n)\\&\le(n-1)V_t^L((n-2)^+)+(n+1)V_t^L(n)+2V_t^L(n-1)+(L-n)(V_t^L(n-1)+V_t^L(n+1)),\end{aligned}$$

betetzen dela frogatzea non $1\le n\le L-1$ den. Eskuineko aldean $2V_t^L(n-1)\le V_t^L((n-2)^+)+V_t^L(n)$ inekuazioa erabiliz, hurrengo inekuazioa lortzen da

$$2nV_t^L(n-1)+2(L-n)V_t^L(n)\le n(V_t^L((n-2)^+)+V_t^L(n))+(L-n)(V_t^L(n-1)+V_t^L(n+1)).$$

Azken hau V_t^L-ren ganbiltasunak inplikatzen du.

Orain min gaiak izango dira aztergai. $n-1$ eta $n+1$ egoeretan har daitezkeen akzioengan jokatuko da frogapenean. t denboran V_t^L ganbila denez, akzio optimoa monotonoa da. Definitu $a_n^*\in\{0,1\}$ n egoerako akzio optimoari, 0 (1) akzioa pasiboa (aktiboa) delarik. Orduan, monotonoa denez, hurrengo hiru aukerak gerta daitezke: $(a_{n-1}^*,a_{n+1}^*)\in\{(0,0),(0,1),(1,1)\}$. Lehenik eta behin suposatu $a^*=(0,1)$. Orduan edozein $1\le n\le L-1$ bada

$$\begin{aligned}&2\min\{-W+\tilde C(n,0)+\omega V_t^L(n),\tilde C(n,1)+\omega V_t^L(n-1)\}\\&\le -W+\tilde C(n,0)+\omega V_t^L(n)+\tilde C(n,1)+\omega V_t^L(n-1)\\&\le -W+\tilde C(n-1,0)+\omega V_t^L(n)+\tilde C(n+1,1)+\omega V_t^L(n-1)\\&=\min\{-W+\tilde C(n-1,0)+\omega V_t^L(n-1),\tilde C(n-1,1)+\omega V_t^L((n-2)^+)\},\\&\quad+\min\{-W+\tilde C(n+1,0)+\omega V_t^L(n+1),\tilde C(n+1,1)+\omega V_t^L(n)\},\end{aligned}\tag{3.7.7}$$

non bigarren inekuazioan C-k eta beraz $\tilde C$-k (3.1.2) betetzen dutela erabili den. $a^*=(1,1)$ den kasuan eta $1\le n\le L-1$

$$\begin{aligned}&2\min\{-W+\tilde C(n,0)+\omega V_t^L(n),\tilde C(n,1)+\omega V_t^L(n-1)\}\\&\le 2\tilde C(n,1)+2\omega V_t^L(n-1)\\&\le \tilde C(n-1,1)+\tilde C(n+1,1)+\omega(V_t^L((n-2)^+)+V_t^L(n))\\&=\min\{-W+\tilde C(n-1,0)+\omega V_t^L(n-1),\tilde C(n-1,1)+\omega V_t^L((n-2)^+)\},\\&\quad+\min\{-W+\tilde C(n+1,0)+\omega V_t^L(n+1),\tilde C(n+1,1)+\omega V_t^L(n)\}.\end{aligned}\tag{3.7.8}$$

Bigarren inekuazioan C (eta beraz $\tilde C$) ganbila dela erabili da eta V_t^L ganbila dela.

$a^*=(0,0)$ bada eta $1\le n\le L-1$ orduan

$$\begin{aligned}&2\min\{-W+\tilde C(n,0)+\omega V_t^L(n),\tilde C(n,1)+\omega V_t^L(n-1)\}\\&\le -2W+2\tilde C(n,0)+2\omega V_t^L(n)\\&\le -2W+\tilde C(n-1,0)+\tilde C(n+1,0)+\omega V_t^L(n-1)+\omega V_t^L(n+1)\\&=\min\{-W+\tilde C(n-1,0)+\omega V_t^L(n-1),\tilde C(n-1,1)+\omega V_t^L((n-2)^+)\},\\&\quad+\min\{-W+\tilde C(n+1,0)+\omega V_t^L(n+1),\tilde C(n+1,1)+\omega V_t^L(n)\}.\end{aligned}\tag{3.7.9}$$

Bigarren inekuazioan C (eta beraz, $\tilde{C}$) ganbila dela eta V_t^L ganbila dela erabili dira.

Beraz, V_{t+1}^L ganbila da. Honek $V_t^L(\cdot)$-ren ganbiltasunaren frogapena amaitzen du. Izan ere, $V_t^L \to V^L$ $t \to \infty$ den heinean [76, 9.4. Kapitulua], eta $V_t^L(\cdot)$-ren ganbiltasunak $V^L(\cdot)$-ren ganbiltasuna inplikatzen du.

3.7.2 3.3. Proposizioaren frogapena

Frogapenean notazioa errazteko k indizearekiko menpekotasuna alde batera utzi da. 2.2. Teoreman proposatu den Whittle indizea kalkulatzeko n eta $n-1$ politika monotonoak hartu behar dira kontutan, non $m \geq n+1$ eta $m \geq n$ egoeretan zerbitzaria aktibo dagoen, hurrenez hurren.

Lehenik eta bein n politika aztertuko da. Izan bedi $f^n(ab)$ eta $f^n(ser)$ sistema utzi duten eta zerbitzatuak izan diren bezeroen frakzioa, hurrenez hurren. Tasen kontserbazio argumentu batek sistemara iritsi diren bezero guztiak sistema utzi dutela edo zerbitzatuak izan direla erakusten du, beraz, $\lambda = \lambda f^n(ab) + \lambda f^n(ser)$. Uzte tasa $\sum_{m=0}^{\infty} \theta m \pi^n(m) + (\theta' - \theta) \sum_{m=n+1}^{\infty} \pi^n(m)$ bidez adieraz daiteke, eta tasakiko erkatuz hurrengo erlazioa lor daiteke

$$\theta \mathbb{E}(N^n) + (\theta' - \theta) \sum_{m=n+1}^{\infty} \pi^n(m) = \lambda f^n(ab) = \lambda(1 - f^n(ser)). \tag{3.7.10}$$

Zerbitzatuak izan direnen tasa $\sum_{m=n+1}^{\infty} \mu \pi^n(m)$ idatz daiteke, eta beraz,

$$\lambda f(ser) = \mu \sum_{m=n+1}^{\infty} \pi^n(m),$$

azken hau (3.7.10) ekuazioan ordezkatuz

$$\mathbb{E}(N^n) = \frac{\lambda}{\theta} + \frac{\theta - \theta' - \mu}{\theta} \sum_{m=n+1}^{\infty} \pi^n(m),$$

lortzen da, non N^n k-klaseko bezeroen kopurua den oreka egoeran n atari-politikapean.

Orain batez besteko mantentze kostua kalkulatuko da. $C_k(n_k, a) = c_k(n_k - a)^+ + c'_k a$ kostua (3.1.4) kostuen erlazioan ordezkatuz $\tilde{C}(n, a) = \tilde{c}n + a(\tilde{c}' - \tilde{c})$ lortzen da, non $\tilde{c}$ eta $\tilde{c}'$ enuntziatuan definitu diren. Batez besteko kostua beraz

$$\begin{aligned}
\mathbb{E}(\tilde{C}(N^n, S^n(N^n))) &= \tilde{c}\mathbb{E}(N^n) + (\tilde{c}' - \tilde{c}) \sum_{m=n+1}^{\infty} \pi^n(m) \\
&= \tilde{c}\frac{\lambda}{\theta} + \left(\frac{\tilde{c}(\theta - \theta' - \mu)}{\theta} + \tilde{c}' - \tilde{c} \right) \sum_{m=n+1}^{\infty} \pi^n(m) = \tilde{c}\frac{\lambda}{\theta} + \left(\tilde{c}' - \frac{\tilde{c}(\theta' + \mu)}{\theta} \right) \sum_{m=n+1}^{\infty} \pi^n(m),
\end{aligned}$$

da. Gai guztiak (2.3.6)-n ordezkatuz

$$W(n) = \frac{\tilde{c}(\mu + \theta')}{\theta} - \tilde{c}', \tag{3.7.11}$$

lortzen da, eta honekin frogapena amaitzen da.

3.7.3 3.4. Proposizioaren frogapena

Frogapenean notazioa errazteko k indizearekiko menpekotasuna alde batera utzi da. $\mu+\theta' = \theta$-ren kasuan indizea (3.2.6)-n lortu da, orduan $\mu+\theta' > \theta$ suposatuko da frogapenean zehar. Lehenik eta behin gogoratu oreka egoerako batez besteko probablititateak $\pi^n(i)$ n politikarentzat eta i egoerarentzat (3.2.3). ekuazioak definitzen dituela. n egoera handietarako Whittle indizea kalkulatzeko, $\pi^n(i) - \pi^{n-1}(i), \forall i \geq 0$ kalkulatu behar da. Izan bedi $i = 0$, hau da,

$$\begin{aligned}\pi^n(0) - \pi^{n-1}(0) =& \frac{\left(\pi^{n-1}(0)\right)^{-1} - \left(\pi^n(0)\right)^{-1}}{\left(\pi^n(0)\pi^{n-1}(0)\right)^{-1}} \\ =& \left(\sum_{i=1}^{\infty}\prod_{m=1}^{i}\frac{q^{n-1}(m-1,m)}{q^{n-1}(m,m-1)} - \sum_{i=1}^{\infty}\prod_{m=1}^{i}\frac{q^{n}(m-1,m)}{q^{n}(m,m-1)}\right)\pi^n(0)\pi^{n-1}(0).\end{aligned}$$

Trantsizio tasen gaineko hurrengo oharra frogapenean zehar erabiliko da:

$$q^n(m,m-1) = q^{n-1}(m,m-1), \;\; \forall\, m \neq n, m \geq 1, \tag{3.7.12}$$

$$q^n(m-1,m) = q^{n-1}(m-1,m), \;\; \forall\, m \geq 1. \tag{3.7.13}$$

Erlazio hauek kontutan hartuz eta $q^n(n,n-1) - q^{n-1}(n,n-1) = \theta - \mu - \theta'$ bada, kakulu batzuen ostean

$$\begin{aligned}\pi^n(0) - \pi^{n-1}(0) = \pi^n(0)\pi^{n-1}(0)\sum_{i=n}^{\infty}\prod_{\substack{m=1\\ m\neq n}}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}\left(\frac{1}{q^{n-1}(n,n-1)} - \frac{1}{q^n(n,n-1)}\right) \\ =\pi^n(0)\pi^{n-1}(0)\frac{\theta-\mu-\theta'}{q^{n-1}(n,n-1)}\sum_{i=n}^{\infty}\prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)},\end{aligned}$$

lortzen da. $q^n(m-1,m) = \lambda$ bada edozein $m \geq 1$, $q^n(m,m-1) = \theta m$ edozein $1 \leq m \leq n-1$ eta $q^n(m,m-1) = \mu + \theta' + \theta(m-1)$ edozein $m \geq n$, $\pi^n(0)$-k (3.2.3) ekuazioko espresioa badu, orduan

$$\frac{\pi^n(0)\pi^{n-1}(0)}{q^{n-1}(n,n-1)} \in \mathcal{O}\left(\frac{1}{n}\right) \qquad \text{eta} \qquad \sum_{i=n}^{\infty}\prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)} \in \mathcal{O}\left(\frac{1}{n!}\right).$$

Bestalde,

$$\pi^n(0) - \pi^{n-1}(0) \in \mathcal{O}\left(\frac{1}{nn!}\right). \tag{3.7.14}$$

Orain $\pi^n(i) - \pi^{n-1}(i)$ kalkula daiteke, edozein $0 < i \leq n-1$. (3.7.13). Ekuazioa erabiliz, eta $i \leq n-1$ bada,

$$\pi^n(i) - \pi^{n-1}(i) = \prod_{m=1}^{i}\frac{q^n(m-1,m)}{q^n(m,m-1)}(\pi^n(0) - \pi^{n-1}(0)),$$

lor daiteke. (3.7.14). Ekuazioa dela-eta, eta $q^n(m, m-1) = \theta m$, $q^n(m-1, m) = \lambda$ edozein $m \leq n-1$, orduan $i \leq n-1$ bada

$$\pi^n(i) - \pi^{n-1}(i) = \frac{\lambda^i}{i!\theta^i}(\pi^n(0) - \pi^{n-1}(0)) \in \mathcal{O}\left(\frac{1}{nn!}\right), \tag{3.7.15}$$

lortzen da. Ordea, $i \geq n$ bada, eta n nahiko handia, orduan

$$\pi^n(i) - \pi^{n-1}(i) = \prod_{m=1}^{i} \frac{q^n(m-1, m)}{q^n(m, m-1)}\pi^n(0) - \prod_{m=1}^{i} \frac{q^{n-1}(m-1, m)}{q^{n-1}(m, m-1)}\left(\pi^n(0) - \pi^n(0) + \pi^{n-1}(0)\right).$$

(3.7.14)-ko oharretik, eta $\prod_{m=1}^{i} \dfrac{q^{n-1}(m-1, m)}{q^{n-1}(m, m-1)} \in \mathcal{O}\left(\frac{1}{i!}\right)$, ekuaziotik

$$\pi^n(i) - \pi^{n-1}(i) = \mathcal{O}\left(\frac{1}{i!n!n}\right) + \prod_{m=1}^{i} \frac{q^n(m-1, m)}{q^n(m, m-1)}\pi^n(0) - \prod_{m=1}^{i} \frac{q^{n-1}(m-1, m)}{q^{n-1}(m, m-1)}\pi^n(0),$$

lortzen da. Kalkulu hainbaten ostean eta (3.7.12) eta (3.7.13) oharretatik

$$\begin{aligned}\pi^n(i) - \pi^{n-1}(i) &= \left(\frac{1}{q^n(n, n-1)} - \frac{1}{q^{n-1}(n, n-1)}\right) \prod_{\substack{m=1\\ m\neq n}}^{i} \frac{q^n(m-1, m)}{q^n(m, m-1)} + \mathcal{O}\left(\frac{1}{i!n!n}\right)\\ &= \frac{\mu + \theta' - \theta}{q^{n-1}(n, n-1)}\pi^n(i) + \mathcal{O}\left(\frac{1}{i!n!n}\right),\end{aligned} \tag{3.7.16}$$

lortzen da $i \geq n$ bada. Gogoratu (3.3.3). Ekuaziotik Whittle indizea $\delta(\mu + \theta') - \delta'\theta' + W^c(n)$ ere idatz daitekeela, non $W^c(n)$ mantentze kostuei dagokien indizea den. $W^c(n)$

$$W^c(n) = \frac{\xi_1(n) + \xi_2(n) + \xi_3(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = \frac{\xi_1(n) + \xi_2(n) + \xi_3(n)}{\pi^n(n) + \mathcal{O}(1/n!n)}, \tag{3.7.17}$$

idatz daiteke, non

$$\begin{aligned}\xi_1(n) &:= \sum_{i=1}^{n-1} C(i, 0)(\pi^n(i) - \pi^{n-1}(i)),\\ \xi_2(n) &:= C(n, 0)\pi^n(n) - C(n, 1)\pi^{n-1}(n),\\ \xi_3(n) &:= \sum_{i=n+1}^{\infty} C(i, 1)(\pi^n(i) - \pi^{n-1}(i)).\end{aligned} \tag{3.7.18}$$

Gogoratu baita $C(n, 1)$ eta $C(n, 0)$ maila finituko polinomioz $P < \infty$ eta $Q < \infty$, goitik bornatuak direla, hurrenez hurren. Beraz, $C(n, a) = E(n, a) + o(1)$ idatz daiteke, n-ren balio handietarako, non $E(n, 1) = \sum_{i=0}^{P} C^{(P,i)}n^i$ eta $C^{(P,i)} := \lim_{n\to\infty} \frac{C(n,1) - \sum_{j=i+1}^{P} C^{(P,j)}n^j}{n^i}$, eta $E(n, 0) = \sum_{i=0}^{Q} E^{(Q,i)}n^i$, eta $E^{(Q,i)} := \lim_{n\to\infty} \frac{C(n,0) - \sum_{j=i+1}^{Q} E^{(Q,j)}n^j}{n^i}$. Orokortasunik galdu gabe suposatu P-k $C^{(P,P)} > 0$ betetzen duela eta Q-k $E^{(Q,Q)} > 0$. Orduan

$$\begin{aligned}
\xi_1(n) &= \sum_{i=1}^{n-1} E(i,0)(\pi^n(i) - \pi^{n-1}(i)) + o(1),\\
\xi_2(n) &= E(n,0)\pi^n(n) - E(n,1)\pi^{n-1}(n) + o(1),\\
\xi_3(n) &= \sum_{i=n+1}^{\infty} E(i,1)(\pi^n(i) - \pi^{n-1}(i)) + o(1).
\end{aligned}$$

Izan bedi $\hat{\xi}_1 := \sum_{i=1}^{n-1} E(i,0)(\pi^n(i) - \pi^{n-1}(i))$, eta (3.7.15). Ekuazioan lortutako emaitzarekin, n-ren balio handietarako $\hat{\xi}_1(n) \in \mathcal{O}\left(\frac{n^{Q-1}}{n!}\right) \subset o(1)$, lortzen da. Beraz, n-ren balio handietarako $\xi_1(n) \in o(1)$. Izan bedi $\hat{\xi}_2(n) := E(n,0)\pi^n(n) - E(n,1)\pi^{n-1}(n)$. (3.7.12) eta (3.7.13) oharrak erabiliz eta hainbat kalkulu egin ostean

$$\hat{\xi}_2(n) = \frac{\prod_{m=1}^{n} q^n(m-1,m)}{\prod_{m=1}^{n-1} q^n(m,m-1)}\left(\frac{E(n,0)\pi^n(0)}{q^n(n,n-1)} - \frac{E(n,1)\pi^{n-1}(0)}{q^{n-1}(n,n-1)}\right),$$

lortzen da. Gogoratu $q^{n-1}(n,n-1) = \mu + \theta' + \theta(n-1)$ eta $q^n(n,n-1) = \theta n$, zeinak (3.7.14)-rekin batera

$$\begin{aligned}
\hat{\xi}_2(n) = &\prod_{m=1}^{n} \frac{q^n(m-1,m)}{q^n(m,m-1)} \frac{\theta n}{q^{n-1}(n,n-1)}\left((E(n,0) - E(n,1))\,\pi^n(0) + \mathcal{O}\left(\frac{n^{P-1}}{n!}\right)\right)\\
&+ \pi^n(n)(\mu + \theta' - \theta)\frac{E(n,0)}{q^{n-1}(n,n-1)},
\end{aligned}$$

inplikatzen duten. n-ren balio handietarako

$$\prod_{m=1}^{n} \frac{q^n(m-1,m)}{q^n(m,m-1)} \frac{\theta n}{q^{n-1}(n,n-1)} \cdot \mathcal{O}\left(\frac{n^{P-1}}{n!}\right) \subset \mathcal{O}\left(\frac{n^{P-1}}{(n!)^2}\right) \subset o(1),$$

denez

$$\xi_2(n) = \frac{\pi^n(n)}{q^{n-1}(n,n-1)}\bigg(\theta n(E(n,0) - E(n,1)) + (\mu + \theta' - \theta)E(n,0)\bigg) + o(1), \tag{3.7.19}$$

ondorioztatzen da. Azkenik, $\hat{\xi}_3(n) := \sum_{i=n+1}^{\infty} E(i,1)(\pi^n(i) - \pi^{n-1}(i))$ kalkulatuko da. (3.7.16) ekuaziotik

$$\hat{\xi}_3(n) = \frac{\mu + \theta' - \theta}{q^{n-1}(n,n-1)} \sum_{i=n+1}^{\infty} E(i,1)\pi^n(i) + \sum_{i=n+1}^{\infty} E(i,1) \cdot \mathcal{O}\left(\frac{1}{i!n!n}\right),$$

ikusten da. n-ren balio handietarako $\sum_{i=n+1}^{\infty} E(i,1) \cdot \mathcal{O}\left(\frac{1}{i!n!n}\right) \subset \mathcal{O}\left(\frac{n^{P-1}}{i!n!}\right) \subset o(1)$, denez

$$\xi_3(n) = \frac{\mu + \theta' - \theta}{q^{n-1}(n,n-1)} \sum_{i=n+1}^{\infty} E(i,1)\pi^n(i) + o(1), \tag{3.7.20}$$

lortzen da. Orain $\xi_1 \in o(1)$ erabiliz, eta (3.7.19)-n eta (3.7.20)-n lortutako $\xi_2(n)$-ren espresioa erabiliz, n handia dela suposatuz, $\frac{\xi_1(n)}{\pi^n(n)} \in o(1)$, ikusten da, eta

$$\begin{aligned}\frac{\xi_2(n)}{\pi^n(n)} &= \frac{\theta n(E(n,0)-E(n,1))}{\mu+\theta'+\theta(n-1)} + \frac{(\mu+\theta'-\theta)E(n,0)}{\mu+\theta'+\theta(n-1)} + o(1)\\ &= E(n,0)-E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta n}E(n,0) + o(1)\\ &= E(n,0)-E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta}\sum_{j=1}^{Q} E^{(P,j)} n^{j-1} + o(1),\end{aligned}$$

eta

$$\begin{aligned}\frac{\xi_3(n)}{\pi^n(n)} =& \frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)} \cdot \sum_{i=n+1}^{\infty} E(i,1) \prod_{m=n+1}^{i} \frac{\lambda}{\mu+\theta'+\theta(m-1)} + o(1)\\ =& \frac{\mu+\theta'-\theta}{\theta n} \sum_{i=n+1}^{\infty}\sum_{j=0}^{P} C^{(P,j)} i^j \left(\frac{\lambda}{\theta m}\right)^{i-n} + o(1).\end{aligned}$$

Izan bedi $\tilde{W}^c(n)$ $W^c(n)$-ren balioa n handia denean. (3.7.17). Ekuazioan $\xi_1(n)/\pi^n(n), \xi_2(n)/\pi^n(n)$-ren eta $\xi_3(n)/\pi^n(n)$-ren espresioak ordezkatuz

$$\begin{aligned}&\tilde{W}^c(n) = (E(n,0)-E(n,1)) + (\mu+\theta'-\theta)/\theta\\ &\times \left(\sum_{j=1}^{Q} E^{(Q,j)} n^{j-1} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j}\left(\frac{\lambda}{\theta}\right)^{j+1}\right) + o(1),\end{aligned}$$

lortzen da $n \to \infty$ den heinean, hau da, (3.3.4)-ko espresioa. $E(n,a)$ ez-beherakorra izateak eta 3.1.2 baldintzak $\tilde{W}^c$ ez-beherakorra izatea inplikatzen dute eta beraz, W^∞ ere ez-beherakorra da, honek frogapena amaitzen du.

3.7.4 3.5. Proposizioaren frogapena

Frogapenean notazioa errazteko k indizearekiko menpekotasuna alde batera utzi da.

$\mu + \theta' = \theta$ kasuan indizea (3.2.6) Ekuazioan lortu da, beraz, $\mu + \theta' > \theta$ suposatuko da frogapenean zehar. Gogoratu (3.3.3) ekuaziotik Whittle indizea $\delta(\mu+\theta') - \delta'\theta' + W^c(n)$ idatz daitekeela, non $W^c(n)$ mantentze-kostuari soilik dagokion. Gogoratu (3.7.17) ekuaziotik $W^c(n)$

$$W^c(n) = \frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))}, \tag{3.7.21}$$

ere idatz daitekeela, non $\xi_i(n)$, $i \in \{1,2,3\}$ (3.7.18) ekuazioak definitzen duen.

Lehenik eta behin $\lim_{\lambda\to 0} \pi^{n-1}(0)/\pi^n(0)$ kalkulatuko da, emaitza hau beranduago erabiliko baita frogapenean. Gogoratu (3.2.3) Ekuazioan definitu diren oreka egoerako probabilitateak. Hau eta (3.7.12) eta (3.7.13) erabiliz

$$\lim_{\lambda\to 0}\frac{\pi^{n-1}(0)}{\pi^n(0)}=\lim_{\lambda\to 0}\frac{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}=1+\lim_{\lambda\to 0}\frac{\sum_{m=n}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}-\sum_{m=n}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}$$
$$=1+\lim_{\lambda\to 0}\frac{\sum_{m=n}^{\infty}\left(\frac{\lambda^m(\mu+\theta'+\theta(n-1))}{\theta n\prod_{i=1}^m q^{n-1}(i,i-1)}-\frac{\lambda^m\theta n}{\theta n\prod_{i=1}^m q^{n-1}(i,i-1)}\right)}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}=1+\frac{(\mu+\theta'-\theta)}{\theta n}\cdot\lim_{\lambda\to 0}\frac{\mathcal{O}(\lambda^n)}{1+\mathcal{O}(\lambda)}=1, \tag{3.7.22}$$

lor daiteke. Emaitza honetatik hurrengoa ondorioztatzen da

$$\lim_{\lambda\to 0}\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)}=\lim_{\lambda\to 0}\frac{\lambda^n/(\theta^n n!)}{-\frac{(\mu+\theta'-\theta)}{\theta n}\left(\frac{\frac{\lambda^n}{(\mu+\theta'+\theta(n-1))\theta^{n-1}(n-1)!}+\mathcal{O}(\lambda^{n+1})}{1+\mathcal{O}(\lambda)}\right)}$$
$$=\lim_{\lambda\to 0}-\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}+o(1)=-\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}. \tag{3.7.23}$$

Orain (3.7.21) Ekuazioko lehenengo gaia kontsideratuko da, hau da,

$$\frac{\sum_{m=0}^{n-1}C(m,0)(\pi^n(m)-\pi^{n-1}(m))}{\sum_{m=0}^{n}\pi^n(m)-\sum_{m=0}^{n-1}\pi^{n-1}(m)}=\frac{\sum_{m=0}^{n-1}C(m,0)\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}(\pi^n(0)-\pi^{n-1}(0))}{\pi^n(n)+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}(\pi^n(0)-\pi^{n-1}(0))}$$
$$=\frac{\sum_{m=0}^{n-1}C(m,0)\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}{\frac{\pi^n(n)}{\pi^n(0)-\pi^{n-1}(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}=\frac{\sum_{m=0}^{n-1}C(m,0)\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}. \tag{3.7.24}$$

non lehenengo inekuazioan (3.7.12) eta (3.7.13) baldintzak erabili diren. (3.7.24) limitea lortzeko $\lambda\to 0$ den heinean, (3.7.23) Ekuazioan lortu den emaitza ordezkatuko da

$$\lim_{\lambda\to 0}\frac{\xi_1(n)}{\sum_{m=0}^{n}\pi^n(m)-\sum_{m=0}^{n-1}\pi^{n-1}(m)}=\lim_{\lambda\to 0}\frac{\sum_{m=0}^{n-1}C(m,0)\frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)}+\sum_{m=0}^{n-1}\prod_{i=1}^m\frac{q^n(i-1,i)}{q^n(i,i-1)}}$$
$$=\lim_{\lambda\to 0}\frac{C(0,0)+\mathcal{O}(\lambda)}{-\frac{\mu+\theta'+\theta(n-1)}{\mu+\theta'-\theta}+1+\mathcal{O}(\lambda)}=-C(0,0)\frac{(\mu+\theta'-\theta)}{\theta n}, \tag{3.7.25}$$

lortzeko.

Orain (3.7.21) Ekuazioko bigarren gaia kontsideratuko da, hau da,

$$\frac{C(n,0)\pi^n(n)-C(n,1)\pi^{n-1}(n)}{\pi^n(n)+\sum_{m=0}^{n-1}\pi^n(n)-\sum_{m=0}^{n-1}\pi^n(n-1)}=\frac{C(n,0)-C(n,1)\frac{\pi^{n-1}(n)}{\pi^n(n)}}{1+\frac{1}{\pi^n(n)}(\pi^n(0)-\pi^{n-1}(0))\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}}$$
$$=\frac{C(n,0)-C(n,1)\frac{\theta n\pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)}}{1+\frac{\theta^n n!}{\lambda^n}(1-\pi^{n-1}(0)/\pi^n(0))\sum_{m=0}^{n-1}\frac{\lambda^m}{\theta^m m!}}. \tag{3.7.26}$$

(3.7.22) Ekuazioan eta (3.7.23) Ekuazioan lortu diren emaitzak (3.7.26) Ekuazioan ordezkatuz

$$
\begin{aligned}
&\lim_{\lambda\to 0}\frac{\xi_2(n)}{\sum_{m=0}^{n}\pi^n(n)-\sum_{m=0}^{n-1}\pi^{n-1}(m)} = \lim_{\lambda\to 0}\frac{C(n,0)-C(n,1)\left(\frac{\theta n}{\mu+\theta'+\theta(n-1)}\right)(1+\mathcal{O}(\lambda^n))}{1-\frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)}(1+\mathcal{O}(\lambda))}\\
&=\lim_{\lambda\to 0}\frac{C(n,0)(\mu+\theta'+\theta(n-1))-C(n,1)\theta n+\mathcal{O}(\lambda^n)}{\theta n(1+\mathcal{O}(\lambda))}\\
&=C(n,0)-C(n,1)+C(n,0)\frac{(\mu+\theta'-\theta)}{\theta n}+\mathcal{O}(\lambda)
\end{aligned}
\tag{3.7.27}
$$

lortzen da. Frogapena bukatzeko (3.7.21) Ekuazioko hirugarren gaia aztertuko da, hau da,

$$
\begin{aligned}
&\frac{\sum_{m=n+1}^{\infty}C(m,1)\pi^n(m)-\sum_{m=n+1}^{\infty}C(m,1)\pi^{n-1}(m)}{\pi^n(n)+\sum_{m=0}^{n-1}\pi^n(m)-\sum_{m=0}^{n-1}\pi^{n-1}(m)}\\
&=\frac{\lambda^n\sum_{m=n+1}^{\infty}\frac{\lambda^{m-n}}{\prod_{i=1}^{n-1}q^n(i,i-1)\prod_{i=n+1}^{m}q^n(i,i-1)}\left(\frac{\pi^n(0)}{q^n(n,n-1)}-\frac{\pi^{n-1}(0)}{q^{n-1}(n,n-1)}\right)}{\lambda^n\left(\frac{\pi^n(0)}{\theta^n n!}+\frac{1}{\lambda^n}\left(\pi^n(0)-\pi^{n-1}(0)\right)\sum_{m=0}^{n-1}\frac{\lambda^m}{m!\theta^m}\right)}\\
&=\frac{\sum_{m=n+1}^{\infty}\frac{\lambda^{m-n}}{\prod_{i=n+1}^{m}q^n(i,i-1)}\left(1-\frac{\theta n\pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)}\right)}{\left(1+\frac{\theta^n n!}{\lambda^n}\left(1-\frac{\pi^{n-1}(0)}{\pi^n(0)}\right)\sum_{m=0}^{n-1}\frac{\lambda^m}{m!\theta^m}\right)}.
\end{aligned}
$$

Azken espresioan (3.7.22) Ekuazioan eta (3.7.23) Ekuazioan lortu diren emaitzak ordezkatuko dira eta

$$
\begin{aligned}
&\lim_{\lambda\to 0}\frac{\xi_3(n)}{\sum_{m=0}^{n}\pi^n(n)-\sum_{m=0}^{n-1}\pi^{n-1}(m)}=\lim_{\lambda\to 0}\frac{\sum_{m=n+1}^{\infty}\frac{\lambda^{m-n}}{\prod_{i=n+1}^{m}q^n(i,i-1)}\left(1-\frac{\theta n\pi^{n-1}(0)}{(\mu+\theta'+\theta(n-1))\pi^n(0)}\right)}{\left(1+\frac{\theta^n n!}{\lambda^n}\left(1-\frac{\pi^{n-1}(0)}{\pi^n(0)}\right)\sum_{m=0}^{n-1}\frac{\lambda^m}{m!\theta^m}\right)}\\
&=\lim_{\lambda\to 0}\frac{\sum_{m=n+1}^{\infty}\frac{\lambda^{m-n}}{\prod_{i=n+1}^{m}q^n(i,i-1)}\left(1-\frac{\theta n}{\mu+\theta'+\theta(n-1)}(1+\mathcal{O}(\lambda^n))\right)}{\left(1-\left(\frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)}+\mathcal{O}(\lambda)\right)\sum_{m=0}^{n-1}\frac{\lambda^m}{m!\theta^m}\right)}=\lim_{\lambda\to 0}\frac{\mathcal{O}(\lambda)}{\frac{\theta n}{\mu+\theta'+\theta(n-1)}+\mathcal{O}(\lambda)}=0,
\end{aligned}
\tag{3.7.28}
$$

dela frogatu. Azkenik (3.7.25) Ekuazioan eta (3.7.27) Ekuazioan eta (3.7.28) Ekuazioan lortu diren emaitzak $\lim_{\lambda\to 0}W^c(n)$-n ordezkatuko dira, eta orduan

$$
\lim_{\lambda\to 0}W^c(n)=C(n,0)-C(n,1)+\frac{(\mu+\theta'-\theta)}{\theta n}(C(n,0)-C(0,0)).
$$

3.7.5 3.6. Proposizioaren frogapena

Frogapenean notazioa errazteko k indizearekiko menpekotasuna alde batera utzi da.

$\mu+\theta'=\theta$ kasuan indizea (3.2.6) Ekuazioan lortu da, beraz, $\mu+\theta'>\theta$ suposatuko da frogapenean zehar. Gogoratu (3.3.3) ekuaziotik Whittle indizea $\delta(\mu+\theta')-\delta'\theta'+W^c(n)$ idatz daitekeela, non $W^c(n)$ mantentze-kostuari soilik dagokion. Gogoratu (3.7.17) ekuaziotik

$$
W^c(n)=\frac{\xi_1(n)+\xi_2(n)+\xi_3(n)}{\pi^n(n)+\sum_{m=0}^{n-1}(\pi^n(m)-\pi^{n-1}(m))},
\tag{3.7.29}
$$

lortzen dela non $\xi_i(n)$, $i \in 1, 2, 3$ (3.7.18) Ekuazioan definitu diren.

Lehenik eta behin $\pi^{n-1}(0)/\pi^n(0)$, kalkulatuko da, zeina frogapenean zehar erabiliko den:

$$\frac{\pi^{n-1}(0)}{\pi^n(0)} = \frac{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} = \left(1 + \frac{\sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^n(i,i-1)} - \sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}\right)$$

$$= 1 + \frac{(\mu + \theta' - \theta)}{\theta n} \cdot \frac{\sum_{m=n}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}}{\sum_{m=0}^{\infty} \frac{\lambda^m}{\prod_{i=1}^m q^{n-1}(i,i-1)}} \tag{3.7.30}$$

$$= 1 + \frac{\mu + \theta' - \theta}{\theta n}(1 + o(1)). \tag{3.7.31}$$

Orain (3.7.29) Ekuazioa kalkulatuko da $\lambda \to \infty$ den heinean. Lehenengo $\xi_1(n)$ gaia aztertuko da. Hainbat kalkuluren ostean

$$\frac{\xi_1(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = \frac{\sum_{m=0}^{n-1} C(m,0)(\pi^n(m) - \pi^{n-1}(m))}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)}$$

$$= \frac{\sum_{m=0}^{n-1} C(m,0)\frac{\lambda^m}{\theta^m m!}}{\frac{\lambda^n/(\theta^n n!)}{1-\pi^{n-1}(0)/\pi^n(0)} + \sum_{m=0}^{n-1} \frac{\lambda^m}{\theta^m m!}}, \tag{3.7.32}$$

azken honetan (3.7.31) Ekuazioa ordezkatuz

$$\frac{\xi_1(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = \mathcal{O}\left(\frac{1}{\lambda}\right), \tag{3.7.33}$$

lortzen da $\lambda \uparrow \infty$ den heinean, edozein n-tarako. Orain (3.7.29) Ekuazioko bigarren gaia kalkulatuko da $\lambda \to \infty$ den heinean. (3.7.31) Ekuazioa erabiliz

$$\lim_{\lambda\to\infty} \frac{\xi_2(n)}{\sum_{m=0}^{n} \pi^n(n) - \sum_{m=0}^{n-1} \pi^{n-1}(m)} = \lim_{\lambda\to\infty} \frac{C(n,0) - C(n,1)\frac{\pi^{n-1}(n)}{\pi^n(n)}}{1 + \frac{\pi^n(0)-\pi^{n-1}(0)}{\pi^n(n)} \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}}$$

$$= \lim_{\lambda\to\infty} \frac{C(n,0) - C(n,1)\frac{\theta n}{\mu+\theta'+\theta(n-1)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}{1 + \frac{1-\pi^{n-1}(0)/\pi^n(0)}{\lambda^n/(\theta^n n!)} \sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}} = C(n,0) - C(n,1), \tag{3.7.34}$$

lortzen da edozein n-tarako. (3.7.29) Ekuazioko hirugarren gaia, hau da,

$$\frac{\xi_3(n)}{\sum_{m=0}^{n} \pi^n(m) - \sum_{m=0}^{n-1} \pi^{n-1}(m)}$$

$$= \frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{n-1} q^n(i,i-1)\prod_{i=n+1}^{m} q^n(i,i-1)}\left(\frac{\pi^n(0)}{\theta n} - \frac{\pi^{n-1}(0)}{\mu+\theta'+\theta(n-1)}\right)}{\pi^n(n) + (\pi^n(0) - \pi^{n-1}(0))\sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}}$$

$$= \frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{m} q^n(i,i-1)}\left(\frac{\theta n}{\mu+\theta'+\theta(n-1)}\left(1 - \frac{\pi^{n-1}(0)}{\pi^n(0)}\right) + \frac{\mu+\theta'-\theta}{\mu+\theta'+\theta(n-1)}\right)}{\lambda^n/(\theta^n n!) + (1 - \pi^{n-1}(0)/\pi^n(0))\sum_{m=0}^{n-1} \frac{\lambda^m}{m!\theta^m}}, \tag{3.7.35}$$

da, non bigarren pausuan $\prod_{i=1}^{n-1} q^n(i,i-1)\prod_{i=n+1}^{m} q^n(i,i-1) = \prod_{i=1}^{m} q^n(i,i-1)/\theta n$ erabili den. Azken ekuazioan (3.7.30) ordezkatu ondoren eta kalkulu batzuen ostean, (3.7.35)

$$\frac{\mu+\theta'-\theta}{\theta n}\frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}}{\frac{\lambda}{\theta n}\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}(1+o(1))},$$

idatz daitekeela ikusten da. Beraz, hirugarren gaia $\lambda\to\infty$ den heinean

$$\frac{\mu+\theta'-\theta}{\theta}\frac{\sum_{m=n+1}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}}{\frac{\lambda}{\theta}\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{i=1}^{m} q^{n-1}(i,i-1)}}+o(1) = \frac{\mu+\theta'-\theta}{\theta}\frac{\sum_{m=n+1}^{\infty} C(m,1)\pi^{n-1}(m)}{\lambda/\theta}+o(1), \tag{3.7.36}$$

idatz daiteke. Azken berdintasuna $\pi^{n-1}(0) = (\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)})^{-1}$ dela-eta betetzen da. Ohartu (3.7.36) Ekuazioa

$$\begin{aligned}&\frac{\mu+\theta'-\theta}{\theta}\left(\frac{\sum_{m=0}^{\infty} C(m,1)\pi^{n-1}(m)}{\lambda/\theta}-\frac{\sum_{m=0}^{n} C(m,1)\pi^{n-1}(m)}{\lambda/\theta}\right)\\ &=\frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta}\left(1-\frac{\sum_{m=0}^{n} C(m,1)\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}{\mathcal{O}(\lambda^n)+\sum_{m=n+1}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}\right),\end{aligned} \tag{3.7.37}$$

idatz daitekeela, non

$$\mathbb{E}(C(N^{n-1},1)) = \frac{\sum_{m=0}^{\infty} C(m,1)\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}{\sum_{m=0}^{\infty}\frac{\lambda^m}{\prod_{j=1}^{m} q^{n-1}(j,j-1)}}.$$

Orduan existitzen da $z\geq 1$ zeinentzat $\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda^z}\to 0$ $\lambda\to\infty$ den heinean, orduan (3.7.37) Ekuazioa

$$\frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta}+o(1),$$

ere idatz daiteke. Beraz, (3.7.29), (3.7.33) eta (3.7.34) Ekuazioetatik

$$W^c(n) = C(n,0)-C(n,1)+\frac{\mu+\theta'-\theta}{\theta}\frac{\mathbb{E}(C(N^{n-1},1))}{\lambda/\theta}+o(1),$$

lortzen da $\lambda\to\infty$ den heinean. Honek frogapena amaitzen du.

3.7.6 3.7. Proposizioaren frogapena

Frogapenean notazioa errazteko k indizearekiko menpekotasuna alde batera utzi da.

$\theta' = \theta$ denez $\mu + \theta' > \theta$ betetzen da. $\delta' = \delta = 0$, $\theta' = \theta$ eta $C(n,a) = C(n)$ direnez $\tilde{C}(n,a) = C(n)$ idatz daiteke. Beraz, hurrengo limitea da aztergai

$$\begin{aligned}\lim_{\theta\to 0} \theta W(n) &= \lim_{\theta\to 0} \frac{\theta \sum_{m=0}^{\infty} C(m) \left(\pi^n(m) - \pi^{n-1}(m)\right)}{\sum_{m=1}^{n-1} (\pi^n(m) - \pi^{n-1}(m)) + \pi^n(n)} \\ &= \varepsilon_1(n)\varepsilon_2(n),\end{aligned}$$

non

$$\varepsilon_1(n) = \lim_{\theta\to 0} \frac{\theta}{\sum_{m=1}^{n-1} (\pi^n(m) - \pi^{n-1}(m)) + \pi^n(n)},$$

eta

$$\varepsilon_2(n) = \lim_{\theta\to 0} \sum_{m=0}^{\infty} C(m) \left(\pi^n(m) - \pi^{n-1}(m)\right).$$

Kontsideratu $\varepsilon_2(n)$. Orduan,

$$\pi^n(0) - \pi^{n-1}(0) \xrightarrow{\theta\to 0} 0.$$

Beraz,

$$\pi^n(m) - \pi^{n-1}(m) \xrightarrow{\theta\to 0} 0, \qquad \forall m < n-1,$$

$$\pi^n(n-1) - \pi^{n-1}(n-1) \xrightarrow{\theta\to 0} (\rho - 1),$$

eta

$$\pi^n(m) - \pi^{n-1}(m) \xrightarrow{\theta\to 0} \rho^{m-n}(1-\rho)^2, \qquad \forall m \geq n.$$

Azken honek

$$\begin{aligned}\varepsilon_2(n) &= -C(n-1)(1-\rho) + \frac{(1-\rho)}{\rho} \sum_{m=n}^{\infty} C(m)(1-\rho)\rho^{m-n+1} \\ &= \frac{(1-\rho)}{\rho}(-C(n-1) + \sum_{m=0}^{\infty} C(m+n-1)(1-\rho)\rho^m),\end{aligned}$$

inplikatzen du.

Hainbat kalkuluren ostean eta $\pi^n(n) \xrightarrow{\theta\to 0} (1-\rho)^{-1}$ dela erabiliz (3.4. Sekzioan aipatu den legez), $\varepsilon_1(n) = 1/\mu$, lortzen da. Honek frogapena amaitzen du.

3.7.7 3.8. Proposizioaren frogapena

Lehenik eta behin k existitzen dela suposatuko da zeinetarako $C_k(0,1) > 0$ betetzen den. Demagun $W = 0$, eta (2.5.1) Ekuaziotik jakina da $\mathcal{C}^{REL}(0) \leq \mathcal{C}^{OPT}$ dela. Demagun baita $\bar{u} \in \mathcal{U}$ politikak aktibo akzioa hartzen duela 0 bezero daudenean sisteman eta, pasiboa dela bestela. Ohartu $\bar{u}$ politikak ez duela inongo *scheduling* erabakirik hartzen. $\mu_k + \theta'_k \geq \theta_k$ denez edozein k-tarako, $\bar{u}$ politikapeko ilararen luzeera estokastikoki goitik bornatua dago edozein $u \in \mathcal{U}$ politikengatik. Ohartu $C_k(0,0) \geq C_k(0,1)$, $\forall k$, hipotesitik, eta (3.1.2) Ekuaziotik edozein n-rako $C_k(n,0) \geq C_k(n,1)$, betetzen dela, zeinak $W_k(n)$ positiboa izatea inplikatzen duten, ikusi 3.2.3. Sekzioa. Beraz, $\mathcal{C}^{WI} \leq \mathcal{C}^{\bar{u}}$ betetzen da. $\frac{\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL}(0)}{\mathcal{C}^{OPT}} \to 0$ betetzen dela erakutsiko da $\lambda \to 0$ den heinean, horrek (2.5.1) Ekuazioarekin batera Whittle indize politika optimoa izatea inplikatzen du.

$W_k(0) = C_k(0,0) - C_k(0,1) \geq 0$, da edozein k-rako. Finkatu $W = 0$, orduan edozein klasetarako $REL(0)$ -1 atariko atari-politika da, hau da, k klasea beti aktibatzen da edozein $n_k > -1$ bada. Beraz, $REL(0)$ politikapean k klasearen oreka egoerako probabilitateak (3.2.3) Ekuazioak definitzen ditu $n = -1$ atari-politikarekin. Orduan

$$\begin{aligned}\mathcal{C}^{REL(0)}(0) &= \sum_{k=1}^{K}\sum_{m=0}^{\infty} C_k(m,1)\pi_k^{-1}(m)\\ &= \sum_{k=1}^{K} C_k(0,1)\pi_k^{-1}(0) + \sum_{k=1}^{K} C_k(1,1)\frac{\lambda\gamma_k}{\mu_k+\theta'_k}\pi_k^{-1}(0) + \mathcal{O}(\lambda^2), \qquad (3.7.38)\end{aligned}$$

betetzen da $\lambda \downarrow 0$ den heinean. Bestalde, $\pi_k^{-1}(0) = (1+\mathcal{O}(\lambda))^{-1}$ da, beraz, $\mathcal{C}^{REL(0)}(0) = \sum_{k=1}^{K} C_k(0,1) + \mathcal{O}(\lambda)$.

$\bar{u} \in \mathcal{U}$ politikapean, edozein k klaseren portaera M/M/∞ ilara bezelakoa da non iritsiera tasak $\lambda\gamma_k$ diren eta irteera tasak $\theta_k n_k$. Orduan $\mathcal{C}^{\bar{u}} = \sum_{k=1}^{K} C_k(0,1)\mathrm{e}^{-\lambda\gamma_k/\theta_k} + \sum_{k=1}^{K}\sum_{m=1}^{\infty} C_k(m,0)\frac{(\lambda\gamma_k)^m}{\theta_k^m m!}\mathrm{e}^{-\lambda\gamma_k/\theta_k} = \sum_{k=1}^{K} C_k(0,1) + \mathcal{O}(\lambda)$, lortzen da.

Beraz,

$$\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}(0) \quad = \quad \mathcal{O}(\lambda). \qquad (3.7.39)$$

Ohartu bestalde, $\lambda \to 0$ limitea hartuz gero, $\mathcal{C}^{OPT} \geq \mathcal{C}^{REL(0)}(0) = \mathcal{O}(1)$ betetzen dela. Azken honek, (2.5.1) eta (3.7.39) ekuazioekin batera

$$\lim_{\lambda\downarrow 0}\frac{\mathcal{C}^{WI}-\mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} \leq \lim_{\lambda\downarrow 0}\frac{\mathcal{C}^{\bar{u}}-\mathcal{C}^{REL(0)}(0)}{\mathcal{C}^{OPT}} = 0,$$

inplikatzen dute.

$C_k(0,1) = 0$ den kasuan, $C_k(0,0) \geq C_k(0,1)$ edozein k klaserako eta $W_k(0) = C_k(0,0) - C_k(0,1) \geq 0$ edozein k-rako. Finkatu $W = 0$, orduan $REL(0)$ politika k klasea aktibatzen duen politika da edozein $n_k \geq 0$ egoeretan. Kontsideratu $\bar{u}$ politika, sisteman dauden bezero kopurua 0 edo 1 denean akzioa aktiboa hartzen duen politika bezala, eta akzio pasiboa bestela. Orduan

$$\mathcal{C}^{\bar{u}} = \sum_{k=1}^{K} C_k(1,1)\frac{\lambda\gamma_k}{\mu_k+\theta'_k}\pi_k^{\bar{u}}(0) + \sum_{k=1}^{K}\sum_{m=2}^{\infty} C_k(m,0)\frac{(\lambda\gamma_k)^m}{(\mu_k+\theta'_k)\theta^{m-1}m!}\pi_k^{\bar{u}}(0),$$

eta $\pi_k^{\bar{u}}(0) = \left(1 + \frac{\lambda\gamma_k}{\mu_k+\theta'_k} + \frac{(\lambda\gamma_k)^2}{(\mu_k+\theta'_k)2\theta_k} + \mathcal{O}(\lambda^3)\right)^{-1}$, $\lambda \to 0$ den heinean.

$\pi_k^{-1}(0) = \pi_k^{\bar{u}}(0) + \mathcal{O}(\lambda^2)$ betetzen da $\lambda \to 0$ limitean. Orduan, $C(1,1)$-i dagokien gaiek $\mathcal{C}^{\bar{u}}$ eta (3.7.38) Ekuazioko $\mathcal{C}^{REL(0)}(0)$-ren espresioetan bat egiten dute $\mathcal{O}(\lambda^2)$ gairaino. Beraz, $\mathcal{C}^{\bar{u}} - \mathcal{C}^{REL(0)}(0) = \mathcal{O}(\lambda^2)$ eta $\mathcal{C}^{OPT} \geq \mathcal{C}^{REL(0)}(0) = \mathcal{O}(\lambda)$, azken hauek behar den soluzioa ondorioztatzen dute, hau da, $\lim_{\lambda\downarrow 0}\frac{\mathcal{C}^{WI}-\mathcal{C}^{OPT}}{\mathcal{C}^{OPT}} = 0$.

3.7.8 3.9. Proposizioaren frogapena

Frogapena erraztearren 2 bezero klase daudela suposatuko da. Tribiala da klase kopuru arbitrario baterako orokortzea. Orokortasunik galdu gabe $\bar{k}=2$ suposatuko da, beraz

$$\lim_{\lambda\to\infty}\frac{w_2}{w_1}>1.$$

3.9. proposizioa frogatzeko hurrengo pausuak jarraituko dira:

- **1. Pausua:** $\bar{W}$ existitzen dela frogatuko da zeinentzat

$$\lim_{\lambda\to\infty}\frac{\mathcal{C}^{REL(\bar{W})}(\bar{W})}{\mathcal{C}^{REL(\bar{W})}}=0$$

  hau da, problema erlaxatuaren soluzioa onargarria da jatorrizko problemarentzat.

- **2. Pausua:** 1. Pausuak $REL(\bar{W})$ politikak $\bar{k}$ klasea soilik zerbitzatuko duela ondorioztatzen da, 1 probabilitatearekin, eta onargarria bilakatzen da jatorrizko problemarentzat. Beraz, 1. Pausuko emaitzak aplikatuz $\lim_{\lambda\to\infty}\frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}}$ lortzen da.

- **3. Pausua** 1. eta 2. Pausuetako emaitzak erabiliz $\lim_{\lambda\to\infty}\frac{\mathcal{C}^{OPT}}{\mathcal{C}^{WI}}$ ondorioztatzen da.

1. Pausua. $W_k(n_k)$ konstantea bada $k=1,2$ orduan enuntziatuko hipotesitik, aurki daiteke $\bar{W}$ konstantea zeinentzat $w_1\le\bar{W}\le w_2$ den. Gogoratu (2.3.2) ekuaziotik

$$\begin{aligned}\mathcal{C}^{REL(\bar{W})}(\bar{W})=&\sum_{k=1}^{2}\mathbb{E}(\tilde{C}(N_k^{REL(\bar{W})},S_k^{REL(\bar{W})}(N_k^{REL(\bar{W})})))\\&-\bar{W}\left(1-2+\sum_{k=1}^{2}\left(1-\mathbb{E}\left(S_k^{REL(\bar{W})}(N_k^{REL(\bar{W})})\right)\right)\right),\end{aligned}$$

betetzen dela. Hipotesiz $w_1<w_2$ denez, $REL(\bar{W})$ politikak ez du inoiz 1 klasea zerbitzatuko eta beti 2 klasea zerbitzatuko du, 2 klaseko bezerorik den heinean. Beraz, hurrengoa froga daiteke

$$\sum_{k=1}^{K}\mathbb{E}(S^{REL(\bar{W})}(N_k^{REL(\bar{W})})=1)\to 1,$$

$\lambda\to\infty$ den heinean, izan ere, lan karga handietan beti aurki daitezke 2 klaseko bezeroak. Hurrengoa

$$\mathcal{C}^{REL(\bar{W})}=\sum_{k=1}^{K}\mathbb{E}(\tilde{C}(N_k^{REL(\bar{W})},S^{REL(\bar{W})}(N_k^{REL(\bar{W})}))),$$

egia denez,

$$\lim_{\lambda\uparrow\infty}\frac{\mathcal{C}^{REL(\bar{W})}(\bar{W})}{\mathcal{C}^{REL(\bar{W})}}=1,$$

betetzen da.

2. Pausua. Orain $\lim_{\lambda\uparrow\infty}\mathcal{C}^{REL(\bar{W})}/\mathcal{C}^{WI}=1$ frogatuko da. Ohartu WI politika eta $REL(\bar{W})$ berdinak direla $N_2>0$ den heinean. $N_2=0$ denean WI politikak 1 klasea zerbitzatzen du, ordez $REL(\bar{W})$

politikak ez du bezerorik zerbitzatzen. $\mu_k+\theta'_k \geq \theta_k$ denez $k=1,2$ denean eta $C_k(m,0) \geq C_k(m,1)$, zuzena da $N_1^{WI} \leq_{st} N_1^{REL}$ dela eta N_2^{WI} eta $N_2^{REL(\bar{W})}$ estokastikoki berdinak direla. Honek zera inplikatzen du $\mathcal{C}^{WI} \leq \mathcal{C}^{REL(\bar{W})}$ dela eta beraz, $1 \leq \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}}$. (2.5.1) Ekuaziotik

$$\mathcal{C}^{REL(\bar{W})}(\bar{W}) \leq \mathcal{C}^{WI} \Longrightarrow \frac{\mathcal{C}^{REL(\bar{W})}(\bar{W})}{\mathcal{C}^{REL(\bar{W})}} \cdot \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} \leq \frac{\mathcal{C}^{WI}}{\mathcal{C}^{WI}} = 1 \Longrightarrow \lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} \leq 1,$$

ondorioztatzen da. Zeinetarako 1. Pausuko emaitza erabili den. Orduan $\lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} \leq 1$ eta $1 \leq \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}}$ diera zeinak $\lim_{\lambda\to\infty} \mathcal{C}^{REL(\bar{W})}/\mathcal{C}^{WI} = 1$, frogatzen duten.

3. Pausua. (2.5.1). Ekuaziotik

$$\lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}(\bar{W})}{\mathcal{C}^{REL(\bar{W})}} \cdot \lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{REL(\bar{W})}}{\mathcal{C}^{WI}} \leq \lim_{\lambda\uparrow\infty} \frac{\mathcal{C}^{OPT}}{\mathcal{C}^{WI}} \leq \frac{\mathcal{C}^{WI}}{\mathcal{C}^{WI}} = 1,$$

lortzen da. 1. eta 2. Pausuetako emaitzak erabiliz $\mathcal{C}^{OPT}/\mathcal{C}^{WI} \to 1$ lortzen da $\lambda \to \infty$ den heinean, zeinak frogapena amaitzen duen.

Atala II

Baliabideen esleipenerako sistema fluidoen kontrol dinamikoa

4. Kapitulua

Indize fluido politika jaiotza-eta-heriotza motako restless bandit*-entzat*

Contents

4.1 Optimizazio problema erlaxatuaren bertsio fluidoa 72

4.2 Indize estokastikoa eta fluidoaren arteko baliokidetasuna trafiko arinean . . 76

4.3 Adibideak 78

4.4 Eranskina 90

I. Atalean jaiotza-eta heriotza motako RBP-entzat errendimendu ona erakusten duten politikak nola lortu azaldu da. Whittle-en indizea efikazki kalkulatu daiteke, 3. Kapituluan uzteak gerta daitezken ilara klase-anitzaren kasurako egin den legez. Halere, orokorrean indizea kalkulatzeko hainbat propietate tekniko frogatu behar dira eta parametroekiko duen menpekotasuna kasu partikularretan bakarrik eskura daiteke. Arazo honi aurre egiteko probema erlaxatuaren bertsio fluidoa proposatuko da. Hurbilketa hau Avram et al. [14] eta Weiss [96] lanek motibatu dute, non optimizazio problema estokastikoak hurbiltzeko kontrol fluido ereduak erabili diren. Hurbilketa honen abantailetako bat, mugatu gabeko optimizazio problema deterministaren soluzio optimoa karakterizatu daitekeela da. Honek indize esplizitu bat lortzea ahalbideratzen du, indize fluido izenez deituko dena. Ondoren, Whittle indize politikaren teoriari dagokion teoria analogoa garatuko da eta indize fluido politika aurkeztu. Azken honek parametroekiko problemak duen dependentzia agertarazten du eta Whittle indizearekin bat dator hainbat kasutan. Whittle-en indize politika eta indize fluido politika hainbat problematan aztertu dira: *e.g.* uzteak gerta daitezkeen zerbitzari-bakar klase-anitzeko ilaran, haririk gabeko sistemetan *scheduling* oportunista egiteko, kontzientzia energetikoz hornitutako zerbitzari-parkeetan, eta inbentario kudeaketan itemak galkorrak direnean. Emaitza numerikoek erakutsi dute indize fluido politika ia optimoa dela.

Kapitulua hurrengo moduan dago egituratuta. 4.1. Sekzioan, 2. Kapituluan aurkeztu den optimizazio problema erlaxatuaren hurbilketa fluidoa proposatuko da. Azken honek indize fluidoa garatzea ahalbideratzen du, eta beraz, indize fluido politika defini daiteke. 4.2. Sekzioan indize fluido politika goian aipatu diren lau problema ezberdinetan aplikatu da. Frogapen gehienak 4.4. Eranskinean aurki daitezke.

4.1 Optimizazio problema erlaxatuaren bertsio fluidoa

Sekzio honetan (2.3.3) mugarik gabeko dimentsio bakarreko optimizazio problemaren bertsio fluidoa ebatziko da, hau da, problema estokastikoaren batez besteko norabidea bakarrik hartuko da kontutan. 4.1.1. Sekzioan dinamika fluidoak deskribatuko dira eta (2.3.3) problemaren bertsio fluidoa. 4.1.2. Sekzioan mugarik gabeko problema fluidoaren soluzioa aurkeztuko da eta indize fluidoa definituko da. 4.1.3. Sekzioan indize fluido politika aurkeztuko da, jatorrizko problemarentzat heuristika gisa balioko duena.

4.1.1 Eredu fluidoa eta *bias*-optimoa izatea

2.3. Sekzioan aurkeztu den mugarik gabeko optimizazio problema estokastikoa hurbilduko da hemen eredu fluido determinista batez, non k *bandit*-ak egoeren espazio jarraia duen, $[0, \infty)$, $\{0, 1, \ldots\}$ espazio diskretoaren ordez. Fluidoen dinamikak prozesu estokastikoaren batez besteko dinamikak soilik kontutan hartuz lortuko dira.

Izan bedi $m_k(t) \in [0, \infty)$ k *bandit*-aren egoera eta $s_k(t) \in \{0, 1\}$ kontrol parametroa. Izan bedi u, $s_k^u(t)$ determinatzen duen kontrola, hau da, k *bandit*-a aktiboa den ala ez.

Hurrengo notazioa erabiliko da a akziopeko norabidea adierazteko:

$$f_k^a(m_k) := b_k^a(m_k) - d_k^a(m_k), \quad a = 0, 1,$$

non $m_k \geq 0$, eta m_k-ren osoak ez diren balioak b_k^0, d_k^0, b_k^1 eta d_k^1 jarraiak izan daitezen definituko dira. Bestalde, $f_k^a(m_k)$ m_k-n ez-gorakorra dela onartuko da $a \in \{0, 1\}$ bada. Dinamika fluidoa u kontrolpean hurrengo moduan idatz daitezke:

$$\frac{\mathrm{d}m_k^u(t)}{\mathrm{d}t} = (1 - s_k^u(t)) f_k^0(m_k^u(t)) + s_k^u(t) f_k^1(m_k^u(t)), \tag{4.1.1}$$

non u kontrolak $m_k^u(t) \geq 0$ izatea bermatzen duen edozein t-rako.

t denboran, (2.3.3) mugarik gabeko problemaren bertsio fluidoaren kostu funtzioa

$$C_k(m_k(t), s_k(t)) := (1 - s_k(t)) C_k(m_k(t), 0) + s_k(t) C(m_k(t), 1),$$

izango da, non m_k-ren osoak ez diren balioak $C_k(m_k, a)$ m_k-n jarraia izan dadin definituko dira. Bestalde, $C_k(m_k, a)$ funtzioa ganbila dela onartuko da edozein m_k eta $a = 0, 1$ badira.

Dinamika fluidoen $(\bar{m}_k, \bar{s}_k)$ oreka puntu bat, $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = 0$ betetzen duen puntu bat izango da, hau da, $(1 - \bar{s}_k) f_k^0(\bar{m}_k) + \bar{s}_k f_k^1(\bar{m}_k) = 0$, non $\bar{s}_k \in [0, 1]$. Hau da, orekan, $\bar{s}_k$ denbora frakzio batean $(1 - \bar{s}_k)$ $a = 1$ $(a = 0)$ akzioa aukeratuko da. Definitu $\bar{s}_k(\bar{m}_k) := f_k^0(\bar{m}_k)/(f_k^0(\bar{m}_k) - f_k^1(\bar{m}_k))$ eta kapitulu honetan zehar $\bar{s}_k(\bar{m}_k)$ $\bar{m}_k$-n hertsiki gorakorra dela onartuko da. Azken hipotesi honen inguruko eztabaida bat 4.2. Oharrean aurki daiteke.

Eredu estokastikoko helburua optimizazio problema erlaxatua ebaztea da *bandit* finko batentzat, hau da, denboran batez besteko kostua minimizatzen da ken lortu den subsidioa, (2.3.3) ekuazioan aipatu den bezala. Orekan, $\bar{s}_k$ batez bestean sistemak aktibo izatea erabaki duen denboraren adierazle da, beraz, (2.3.3)-ren bertsio fluidoa $EC_k(\bar{s}, W)$ orekako kostua minimizatzea izango da, non

$$EC_k(\bar{s}, W) := (1 - \bar{s}_k) C_k(\bar{m}_k, 0) + \bar{s}_k C(\bar{m}_k, 1) - W(1 - \bar{s}_k).$$

Oreka puntu optimoa (m_k^*, s_k^*) izendatuko da eta beraz, oreka egoerako kostu optimoa W subsidioa emanda

$$EC_k^*(W) := (1 - s_k^*)(C_k(m_k^*, 0) - W) + s_k^* C_k(m_k^*, 1), \tag{4.1.2}$$

da.

Denboran batez besteko kostu optimoa hainbat kontrolek lortuko dute, m_k^* orekatzat duen edozein kontrolek, orduan *bias*-optimoak diren kontrolak izango dira aztergai. Hau da, oreka optimoa lortzen duten kontrol guztien artetik, orekara iristeko kostua minimizatzen duten kontrolak (*bias*-optimoak). Beraz, helburua *bias*-kostua minimizatzen duten u kontrolak aurkitzea izango da, hau da,

$$J_k^u(m_k(0), W) := \int_0^\infty \big(C_k(m_k(t), s_k^u(t)) - W(1 - s_k^u(t)) - EC_k^*(W)\big)\mathrm{d}t. \tag{4.1.3}$$

Bestalde, $J_k(m_k(0), W) := \min_u J_k^u(m_k(0), W)$ definituko da.

Kontrol optimoaren teoriak erakusten du, kontrol bat *bias*-optimoa izan dadin baldintza nahikoa HJB Ekuazioa ebaztea dela, ikusi Section 1.3.4:

$$EC_k^*(W) = \min\Big(C_k(m_k, 1) + f_k^1(m_k)\frac{\partial J_k(m_k, W)}{\partial m_k}, C_k(m_k, 0) - W + f_k^0(m_k)\frac{\partial J_k(m_k, W)}{\partial m_k}\Big). \tag{4.1.4}$$

Orduan, m_k egoera finko batentzat, egoera horretako akzio optimoa (4.1.4) ekuazioaren eskuinaldea minimizatzen duen akzioak definitzen du.

Hurbilketa honen abaitailarik garrantzitsuena (4.1.4) orokorrean ebatzi daitekela da, ikusi 4.1. Proposizioa, baina (2.3.3) ebazteko (baliokideki (2.3.4)) problema erlaxatuaren soluzio optimoak atari egitura duela frogatu behar da.

4.1 Oharra. (4.1.3) *lortzeko modu alternatibo bat kostu deskontatua kontsideratzea da*

$$\mathcal{C}^\phi(\beta) := \sum_{k=1}^{K} \mathbb{E}\left(\int_0^\infty \mathrm{e}^{-\beta t} C_k(N_k^\phi(t), S_k^\phi(\vec{N}^\phi(t)))\ dt\right),$$

non $\beta > 0$ deskontu faktorea den, eta gero bere bertsio fluidoa hartzea. Horrek kontrol problema determinista bat lortzea ahalbideratzen du kostu deskontatuarekin eta hau orokorrean problema zaila da. 2.3. Sekzioan egin den bezela, zerbitzuen gaineko baldintzak erlaxatu daitezke eta bandit aktiboen kopurua $M/(1-\beta)$ balioaz bornatzen da. Bandit bakar baterako helburua mugarik gabeko problema fluidoa kostu deskontatuarekin ebaztea da, hau da, u kontrola aurkitzea zeinak $J_k^{u,\beta}(m_k(0), W) := \int_0^\infty \mathrm{e}^{-\beta t}(\big(C_k(m_k(t), s_k^u(t)) - W(1 - s_k^u(t))\big))dt$, minimizatzen duen. Beraz, mugarik gabeko problema fluidoaren, kostu deskontatuak onartuz, soluzioa

$$\begin{aligned}\beta J_k^\beta(m_k, W) = \min_s(&C_k(m_k, 1) + \beta f_k^1(m_k)\partial J_k^\beta(m_k, W)/\partial m_k,\\ &C_k(m_k, 0) - W + \beta f_k^0(m_k)\partial J_k^\beta(m_k, W)/\partial m_k),\end{aligned} \tag{4.1.5}$$

ebatziz lortzen da, ikusi [76, 10. Kapitulua], non $J_k^\beta(m_k, W) = \min_u J_k^{u,\beta}(m_k, W)$. Ohartu $\beta \to 1$ den heinean $\beta J_k^\beta(m_k, W) \to EC_k^*(W)$, ikusi [76, 8.2.5. Korolarioa], eta beraz (4.1.5) ekuazioak (4.1.4)-ra konbergitzen du.

4.2 Oharra. *3.4. Sekzioan azpimarratu den legez $\sum_{m=0}^{n} \pi_k^n(m)$ hertsiki gorakorra izateak (Whittle indizearen indize gaitasunerako beharrezkoa dena) hainbat eredu uzten ditu analisitik at. Testuinguru fluidoan $\bar{s}_k(\bar{m}_k)$ $\bar{m}_k$-n hertsiki gorakorra izateak ere hainbat eredu uzten ditu analisitik kanpo. Izan bedi M/M/1 ilara bat iritsierak kontrolatuak direnean, eta iritsierak banaketa esponentziala jarraitzen dute $\mu_k a$ tasarekin, non $a = 0,1$ eta irteerak Poisson prozesu bat jarraitzen dute $\mu_k < \lambda_k$ tasarekin. Orduan, $(1 - \bar{s}_k)f_k^0(m_k) + \bar{s}_k f_k^1(m_k) = -\mu_k + \bar{s}_k\lambda_k$. Azken honek 0 ematen du $\bar{s}_k = \mu_k/\lambda_k$ denean. Beraz, $\bar{s}_k = \mu_k/\lambda_k$ denean edozein m_k oreka puntua da. Azken honek sistemak pasibo izatea erabakitzen duen denbora frakzioa ez dela aldatzen oreka puntuarekiko adierazten du. Beraz, pasibo izateko subsidioak ez du egoeren arteko bereizketarik egiten. 4.3. Sekzioan arazo honi irtenbidea ematen zaio itemak galkorrak direla kontsideratuz inbentario problema batean.*

4.1.2 Kontrol optimo fluidoa eta indize fluidoa

Sekzio honetan (4.1.3) mugarik gabeko problema fluidoaren soluzio optimoa ebatziko da *bandit* batentzat. Soluzio hau indize fluido funtzio batek deskribatzen du, zeinak espresio esplizitu bat duen. 4.1.3. Sekzioan definitu den indize fluidoan oinarrituz jatorrizko eredu estokastikorako heuristika bat definitu da, zeina ia optimoa dela erakutsiko den 4.3. Sekzioan.

Indize fluidoa definitzeko, hurrengo notazioa behar da: m_k^a $f_k^a(m_k) = 0$ betetzen duen m_k-ren balioa da, $a = 0,1$ bada. Bestalde, $m_k^a = \infty$ definituko da $f_k^a(m_k) > 0$ bada edozein $m_k \geq 0$-rentzat, eta $m_k^a = 0$ $f_k^a(m_k) < 0$ bada edozein $m_k \geq 0$, hau da, $m_k^a \in [0, \infty)$. Indize fluidoaren egiturak m_k^1 eta m_k^0-ren ordenarekiko dependentzia du. 4.1. Irudian $m_k^1 < m_k^0$ den kasurako norabideak erakutsi dira. Indize fluidoaren forma m_k-k hartzen duen balioarekikoa da, hau da $m_k < m_k^1$ den, $m_k \in [m_k^1, m_k^0]$ den, edo $m_k > m_k^0$ den araberakoa. Lehenengo kasuan, bi norabideak $f_k^0(m_k)$ eta $f_k^1(m_k)$ positiboak dira, bigarren kasuan norabideak bidirekziozkoak dira, eta hirugarren kasuan norabideak negatiboak dira.

$f_k^a(\cdot)$ funtzioa ez-gorakorra dela onartu da $a = 0,1$ denan eta $\bar{s}_k(\bar{m}_k)$ hertsiki monotonoa dela $\bar{m}_k$-n. Indize fluido politika definitzeko, hurrengo definizioa eta hipotesiak beharko dira.

4.1 Definizioa. *$m_k^0 > m_k^1$ bada, finkatu $\bar{a} = 1$ eta $m_k^1 \geq m_k^0$ bada finkatu$\bar{a} = 0$.*

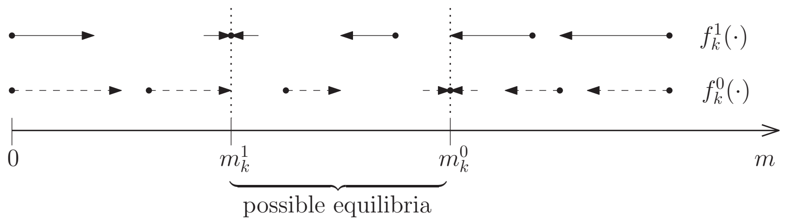
4.1 Hipotesia. *Hurrengo hipotesiak egingo dira:*

- *$f_k^a(m_k)$ diferentziagarria da $[m_k^{\bar{a}}, m_k^{1-\bar{a}}]$ tartean $a = 0,1$ bada.*
- *$f_k^a(m_k)$ ganbila da m_k $a = 0,1$ bada.*
- *$f_k^{\bar{a}}(m_k) - f_k^{\bar{a}}(\bar{m}_k) \geq (\leq) f_k^{1-\bar{a}}(m_k) - f_k^{1-\bar{a}}(\bar{m}_k)$, edozein $m_k \leq (\geq)\bar{m}_k$-rako, non $\bar{m}_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}]$.*
- *$1 - \bar{a} + (2\bar{a} - 1)\bar{s}_k(\bar{m}_k)$ ganbila da $\bar{m}_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}]$ tartean.*

4.1. Hipotesiko item-ak erraz ikus daiteke betetzen diren ala ez problema partikularrentzat. Hau egin da 4.3. Sekzioko lau adibideentzat.

Hurrego proposizioan indize fluidoaren espresioa aurkeztuko da eta (4.1.3) problema fluidoaren soluzioa enuntziatu. Frogapena Eranskinean aurki daiteke eta 4.2. Leman oinarritzen da.

4.1 Proposizioa. *Demagun 4.1. Hipotesiak betetzen direla. Izan bedi $C_k(m_k, a)$, $a = 0,1$, diferentziagarria $m_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}]$ tartean, ganbila eta ez-beherakorra m_k-n eta $a = 0,1$ bada*



Irudia 4.1: Oreka eta norabideen ilustrazioa $m_k^1 < m_k^0$ den kasuan.

Demagun $C_k(m_k, \bar{a}) - C_k(\bar{m}_k, \bar{a}) \leq C_k(m_k, 1-\bar{a}) - C_k(\bar{m}_k, 1-\bar{a})$. *Definitu*

$$w_k^{(a)}(m_k) := (f_k^1(m_k) - f_k^0(m_k))\frac{C_k(m_k, a) - C_k(m_k^a, a)}{f_k^a(m_k)}, a = 0, 1,$$
$$w_k^{(2)}(m_k) := \frac{(f_k^1(m_k) - f_k^0(m_k))(f_k^0(m_k)\frac{\mathrm{d}C_k(m_k,1)}{\mathrm{d}m_k} - f_k^1(m_k)\frac{\mathrm{d}C_k(m_k,0)}{\mathrm{d}m_k})}{f_k^0(m_k)\frac{\mathrm{d}f_k^1(m_k)}{\mathrm{d}m_k} - f_k^1(m_k)\frac{\mathrm{d}f_k^0(m_k)}{\mathrm{d}m_k}},$$

eta demagun $(2\bar{a}-1)w_k^{(i)}(m_k)$, $i = 0, 1, 2$, *ez-beherakorra dela* m_k*-rako. Orduan* (4.1.3)*-ren soluzio optimoa* $s_k(t) = 1$ *da* $W \leq w_k(m_k)$ *bada eta* $s_k(t) = 0$ $W > w_k(m_k)$ *bada, non* $w_k(m_k)$ *funtzio jarraia den eta*

$$w_k(m_k) := C_k(m_k, 0) - C_k(m_k, 1) + \begin{cases} w_k^{(\bar{a})}(m_k) & \textit{if } m_k < m_k^{\bar{a}}, \\ w_k^{(2)}(m_k) & \textit{if } m_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}], \\ w_k^{(1-\bar{a})}(m_k) & \textit{if } m_k > m_k^{1-\bar{a}}, \end{cases}$$

den.

$w_k(\cdot)$ funtzioari *indize fluidoa* deituko zaio.

Indize fluidoak ondorengo interpretazioa du. $EC_k(\bar{s}_k, W)$ funtzioa ganbila da $\bar{s}_k$-n eta $\partial EC_k(\bar{s}_k, W)/\partial \bar{s}_k = 0$ betetzen da baldin eta soilik baldin $W = C_k(m_k, 0) - C_k(m_k, 1) + w_k^{(2)}(m_k)$. Hau da, $m_k \in [m_k^{\bar{a}}, m_k^{1-\bar{a}}]$ denean indize fluidoa orekan kostua minimizatzen duen W-ren balioa da. Kalkulu hauek 4.1. Proposizioaren frogapenean aurki daitezke. Gainera, $m_k < m_k^{\bar{a}}$ edo $m_k > m_k^{1-\bar{a}}$ direnean, indize fluidoa $w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + w_k^{(a)}(m_k)$ hurrengo moduan idatz daiteke:

$$w_k(m_k) = f_k^{1-a}(m_k)\left(\frac{C_k(m_k, 0) - C_k(m_k^a, a)}{f_k^0(m_k)} - \frac{C_k(m_k, 1) - C_k(m_k^a, a)}{f_k^1(m_k)}\right),$$

$a = 0, 1$ bada. Azken hau (3.3.2). Ekuazioaren bertsio fluido bezela uler daiteke eta antzeko interpretazioa du.

4.1. Proposiziotik ikus daiteke $w_k(m)$ monotonoa izateak (4.1.3) problemaren soluzio optimoa atari motako izatea inplikatzen duela: $m_k^1 < m_k^0$ ($m_k^1 > m_k^0$) den kasuan, $w_k(\cdot)$ ez-beherakorra (ez-gorakorra) izateak 0-1 (1-0) motako atari-politika bat optimo izatea inplikatzen du, hau da, pasibo akzioa optimoa da baldin eta soilik baldin $m_k \leq m_k'(W)$ ($m_k \geq m_k'(W)$), non $m_k'(W)$ balioak $w_k(m_k'(W)) = W$ betetzen duen. Eredu estokastikoan atari-politikak optimo izatea kasuz kasu frogatu beharreko propietatea da.

$w_k(m)$ monotonoa izatea erraz ziurtatu daitekeen propietatea da. Honek eredu estokastikoarekiko abantaila bat suposatzen du, izan ere atari-politikak optimoak direla jaiotza-eta-heriotza motako prozesuentzat zaila gerta daiteke, baita indize gaitasuna frogatzea ere. 4.3. Sekzioan indize fluidoa monotonoa dela ikusiko da lau adibideentzat. Hurrengo lemak baldintza nahikoak eskeintzen ditu $w_k(\cdot)$ monotonoa izan dadin.

4.1 Lema. *Izan bedi $C_k(m_k, 1) = C_k(m_k, 0)$ eta $\frac{\mathrm{d}f_k^1(m_k)}{\mathrm{d}m_k} = \frac{\mathrm{d}f_k^0(m_k)}{\mathrm{d}m_k}$. Demagun $C_k(m_k, 1)$ ez-beherakorra dela m_k-n eta demagun $C_k(m_k, 1)$ eta $f_k^1(m_k)$ polinomioak direla $P > 0$ eta $\alpha \geq 0$ mailakoak, hurrenez hurren. Orduan, $(2\bar{a} - 1)w_k(m_k)$ ez-beherakorra da $f_k^{\bar{a}}(m_k) - f_k^{1-\bar{a}}(m_k) < 0$ bada.*

Frogapena. Frogapena $C_k(m_k, 1) = C_k(m_k, 0)$ eta $\frac{\mathrm{d}f_k^1(m_k)}{\mathrm{d}m_k} = \frac{\mathrm{d}f_k^0(m_k)}{\mathrm{d}m_k}$ 4.1. Proposizioko espresioetan ordezkatuz lortzen da, $f_k^a(\cdot)$ ez-gorakorra dela erabiliz $a = 0, 1$ bada. □

2.3. Sekzioan indize gaitasuna izatearen propietatea definitu da, zeinak jatorrizko problemarentzat heuristika bat definitzea ahalbideratzen duen. Eredu fluidorako definizio bera erabiliko da, hau da, *bandit* fluidoak indize gaitasuna du W handitzen den heinean pasibo akzioa optimoa den egoeren multzoa handitzen bada. Propietate hau berehalakoa da $D_k(W) = \{m_k : W > w_k(m_k)\}$ den heinean, ikusi 2.1. Definizioa. Eredu estokastikorako indize gaitasuna kasuz kasu aztertu beharreko propietatea da, hemen ez bezala.

4.3 Oharra. *Hurbilketa honen orokortasuna hurrengoak erakusten du, problema klasikoei aplikatzen zaienean ezagunak diren indize politikak berreskuratzen dira. Adibidez, frogatu daiteke ilara klase-anitz batean mantentze-kostuak linealak badira indize fluidoa $c\mu$ bilakatzen da, eta mantentze-kostuak ganbilak direnean $c\mu$-Orokortuarekin bat dator (zeina [69]-n aurkeztu zen eta trafiko geldoan optimoa dela ikusi zen). Uzteak gerta daitezken ilara klase-anitz batean mantentze-kostuak linealak badira indizea $c\mu/\theta$ indizearekin bat dator (zeina [8] artikuluan proposatu zen eta asintotikoki optimoa dela ikusi zen).*

4.1.3 Indize fluido politika

Indize gaitasunak (3.1.3) problemarentzat heuristika bat definitzea ahalbideratzen du, $w_k(\cdot)$ indize fluidoan oinarritzen dena.

4.2 Definizioa (Indize fluido politika)**.** *Demagun t denboran $\vec{N}(t) = \vec{n}$ egoeran dagoela sistema. Indize fluido politikak $w_k(n_k)$ indize ez negatibo eta handienak dituzten M bandit-ak zerbitzatzea agintzen du.*

4.3. Sekzioan emaitza numerikoak aurkeztuko dira indize fluido politikaren errendimendua ia optimoa dela erakusten dutenak. Bestalde, numerikoki erkatu dira indize fluidoa eta Whittle indizea eredu estokastikoan.

4.2 Indize estokastikoa eta fluidoaren arteko baliokidetasuna trafiko arinean

Sekzio honetan Whittle indizea 2.1. Korolarioak definitzen duela onartuko da eta indize fluidoa 4.1. Proposizioak. Lehenik eta behin trafikoa arina denean bat datozela frogatuko da. Bigarrenik, egoerak balio handiak hartzen dituenean Whittle indizea eta indize fluidoa bat datozela ikusiko da uzteak gerta daitezken ilara klase-anitz baterako (3. Kapituluko eredurako).

Gogoratu jaiotza tasak $b_k^a(\cdot)$-k definitzen dituela $a = 0, 1$ denean. Demagun lehenik eta behin $b_k^a(m_k) = \lambda\gamma_k$ edozein m_k-rako, 0-1 atari-politikak optimoan diren problemetan, eta $b_k^a(m_k) = \lambda\gamma_k a$ edozein m_k-rako, atari-politika optimoak 1-0 motakoak direnean. Hemen $\sum_{k=1}^K \gamma_k = 1$ onartuko da, non λ iritsiera tasen intentsitatea den. Trafiko arina kontsideratuko da, hau da, iritsiera tasa 0-runtz doanean $\lambda \to 0$.

Lehenik eta behin, 4.2. Proposizioan, Whittle indizea kalkulatuko da trafikoa arina denean. Frogapena 4.4.3. Eranskinean aurki daiteke.

4.2 Proposizioa. *Izan bedi $W_k(\cdot)$ (2.3.6) ekuazioak definituta. Orduan $W_k(n_k) = W_k^{LT}(n_k) + o(1)$ $\lambda \downarrow 0$ den heinean, non*

$$W_k^{LT}(n_k) := C_k(n_k, 0) - C_k(n_k, 1) + (d_k^1(n_k) - d_k^0(n_k))\frac{C_k(n_k, \bar{a}) - C_k(0, \bar{a})}{d_k^{\bar{a}}(n_k)},$$

eta $\bar{a} = 0$ (2.3.3) ebazten duen politika 0-1 atari motakoa bada eta $\bar{a} = 1$ 1-0 atari motakoa bada.

Hipotesiz, $d_k^a(m_k) = \lambda\gamma_k > 0$ edozein $m_k > 0$ bada eta $a = 0, 1$ bada, honek $m_k^0 \to 0$ eta $m_k^1 \to 0$ $\lambda \to 0$ doan heinean inplikatzen ditu. Beraz, indize fluidoa trafikoa arina denean $w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + w_k^{(\bar{a})}(m_k)$ da. Hurrengo proposizioan Whittle indizea eta indize fluidoaren arteko baliokidetasuna ezarriko da trafikoa arina denean. Frogapena 4.4.4. Eranskinean aurki daiteke.

4.3 Proposizioa. *Demagun $d_k^a(m_k) > 0$ edozein $m_k > 0$-rako $a = 0, 1$ bada. Izan bedi $W_k(\cdot)$ (2.3.6) ekuazioak definituta. Demagun (4.1.3) problemaren soluzio optimoa n_k atari-politika dela. Orduan*

$$\lim_{\lambda\downarrow 0} W_k(m_k) = \lim_{\lambda\downarrow 0} w_k(m_k).$$

Orain 3. Kapituluan aztergai izan den uzteak gerta daitezkeen ilara klase-anitza, hau da, $b_k^a(m_k) = \lambda_k$ eta $d_k^a(m_k) := (\mu_k + \theta_k' - \theta_k)a + \theta_k m_k$. Hurrengo proposizioan $w_k(n_k)$ indize fluidoa (2.3.6) ekuazioan emandako Whittle indizearekin bat datorrela frogatuko da egoerak balio handiak hartzen dituenean, eta $C_k(m, 1)$ eta $C_k(m, 0)$ kostu funtzioak maila finituko polinomioez, $P_k < \infty$ eta $Q_k < \infty$ mailakoak, bornatuak daudenean, hurrenez hurren. Beraz, $C_k(n_k, a) = E_k(n_k, a) + o(1)$ idatz daiteke n_k-ren balio handietarako, non $E_k(n_k, 1) = \sum_{i=0}^{P_k} C_k^{(P_k,i)} n_k^i$ eta

$$C_k^{(P_k,i)} := \lim_{n_k\to\infty} \frac{C_k(n_k, 1) - \sum_{j=i+1}^{P_k} C_k^{(P_k,j)} n_k^j}{n_k^i},$$

eta $E_k(n_k, 0) = \sum_{i=0}^{Q_k} E_k^{(Q_k,i)} n_k^i$,

$$E_k^{(Q_k,i)} := \lim_{n_k\to\infty} \frac{C_k(n_k, 0) - \sum_{j=i+1}^{Q_k} E_k^{(Q_k,j)} n_k^j}{n_k^i}.$$

4.4 Proposizioa. *Demagun $C_k(n_k, 1)$ eta $C_k(n_k, 0)$ P_k eta Q_k mailako polinomioez goitik bornaturik daudela, hurrenez hurren. Orduan, uzteak gerta daitezkeen eredurako*

$$\lim_{n_k\to\infty} \frac{W_k(n_k)}{w_k(n_k)} = 1. \tag{4.2.1}$$

Gainera $P_k = Q_k$ *bada eta* $C_k^{(P_k,i)} = E_k^{(P_k,i)}$ *edozein* $i \in \{2, \ldots, P_k\}$*, orduan* $n_k \to \infty$ *den heinean*

$$W_k(n_k) = w_k(n_k) + o(1). \tag{4.2.2}$$

Adibide gisa kontsideratu $C_k(n_k, a) = C_k(n_k)$ edo $C_k(n_k, a) = C_k((n_k - a)^+)$. Orduan $Q_k = P_k$, eta beraz (4.2.1) betetzen da. $C_k(n_k, a) = C_k(n_k)$ bada, orduan gainera (4.2.2) betetzen da. 4.4. Proposizioaren frogapena 4.4.5. Eranskinean aurki daiteke.

4.3 Adibideak

Sekzio honetan indize politika estokastikoa eta fluidoa aztertuko dira lau aplikazio adibide ezberdinetan, zeintzuk jaiotza-eta-heriotza motako prozesuak jarraitzen dituzten. Helburua hurrengo lau problemetan erabaki bat hartzen laguntzea da: (i) uzteak gerta daitezken ilara zerbitzari-bakar klase anitz batean *scheduling*-a, haririk gabeko sarean *scheduling* oportunista, azken bi hauek 2.1. Irudiko (ezkerraldeko) dagokion problema motei dagokie, (iii) kontzientzia energetikoz hornitutako zerbitzari-parke batean blokeatzea/bideratzea, eta (iv) inbentarioen kudeaketa itemak galkorrak direnean, azken bi hauek 2.1. Irudiko (eskuinaldeko) problema motei dagokie. Kasu guzti horietan, Whittle indizearen eta indize fluido politiken errendimendua erkatu da soluzio optimoarekiko. Soluzio optimoa lortzeko *value iteration* algoritmoa erabili da, ikusi [76] eta 1.3.3. Sekzioa teknika honi buruzko eztabaida baterako. Ondorio nagusia Whittle indize eta indize fluido politikek errendimendu ia optimoa erakusten dutela da.

4.3.1 Uzteak gerta daitezkeen ilara klase-anitzean *scheduling*-a

Sekzio honetan 3. Kapituluan aztergai izan den uzteak gerta daitezkeen ilara klase-anitza kontsideratu da. k klaseko bezeroak Poisson prozesu bat jarraitzen dute λ_k tasa duena eta zerbitzuak banaketa esponentziala jarraitzen du $1/\mu_k$ batez besteko balioarekin. Ilaran zain dauden bezeroak banaketa esponentziala jarraitzen duen eta $1/\theta_k$ batez bestekoa duen denbora baten ostean uzten dute sistema, eta zerbitzua jasotzen ari diren bezeroek banaketa esponentziala jarraitzen duen eta $1/\theta'_k$ batez bestekoa duen denbora baten ostean uzten dute sistema. MEP-ren transizio tasak $b_k^a(n_k) = \lambda_k$, and $d_k^a(n_k) = (\mu_k + \theta'_k)a + \theta_k(n_k - a)^+$ dira $a = 0, 1$ bada. Gogoratu $\mu_k + \theta'_k \geq \theta_k$ hipotesia ere.

Helburua batez besteko mantentze kostua minimizatzea da, (3.1.4)-en definitu den bezela, $\tilde{C}_k(N_k, a)$ mantentze-kostuen eta uzte-kostuen arteko batura da N_k k klaseko bezero daudenean sisteman.

Testuinguru honetan problema erlaxatuaren soluzio optimoa atari motakoa da 0-1 egiturarekin. Azken hau 3.1. Proposizioan frogatu da.

k klaseko bezeroen oreka egoerako probabilitateak n_k atari-politikapean (3.2.3) Ekuazioak definitzen ditu. Espresio hauek 3.2. Proposizioan erabili dira indize gaitasuna frogatzeko eredu honentzat. Beraz, 2.1. Korolariotik Whittle indizea (2.3.6) ekuazioak definitzen duela ondoriozta daiteke (2.3.6) ez-beherakorra denean. 3. Kapituluan hainbat propietate garatu dira eredu honen Whittle indizearentzat. Orain helburua indize fluidoa garatzea da, erabilgarria den espresio esplizitu bat izatea ahalbideratzen duena edozein kostu funtziorako.

Fluidoaren dinamika

$$\begin{aligned}\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} =&\lambda_k - s_k(t)(\mu_k + \theta'_k + \theta_k(m_k(t)-1)) - (1-s_k(t))\theta_k m_k(t)\\ =&\lambda_k - (\mu_k + \theta'_k - \theta_k)s_k(t) - \theta_k m_k(t),\end{aligned}$$

da, non $s_k(t) \in \{0,1\}$. Beraz, $m_k^0 = \lambda_k/\mu_k$ eta $m_k^1 = (\lambda_k - (\mu_k + \theta'_k - \theta_k))/\theta_k$, hau da, $\bar{m}_k$ oreka puntuak $\bar{m}_k \in [\max(0, m_k^1), \lambda_k/\theta_k]$ betetzen dute. 4.1. Proposiziotik indize fluidoa garatu daiteke, zeinak (4.1.3) problema minimizatzen duen politika deskribatzen duen. Lehenik eta behin indize fluidoa aurkeztuko da eta gero indize fluido politika optimoa dela frogatuko da, ikusi 4.5. eta 4.6. Proposizioak, hurrenez hurren.

4.5 Proposizioa. *Demagun $C_k(m_k, a)$ diferentziagarria dela, ganbila eta ez-beherakorra m_k−n. Bestalde, onartu $C_k(m_k, 0) - C_k(m_k, 1)$ ganbila eta ez-beherakorra dela m_k-n. Orduan indize fluidoa ez-beherakorra da eta*

$$\begin{aligned}w_k(m_k) :=&C_k(m_k,0) - C_k(m_k,1) + \delta_k(\mu_k + \theta'_k) - \delta'_k\theta'_k\\ &+\begin{cases} w_k^{(1)}(m_k) & \textit{if } 0 \le m_k < \max\left(0, \frac{\lambda_k - (\mu_k + \theta'_k - \theta_k)}{\theta_k}\right),\\ w_k^{(2)}(m_k) & \textit{if } \max\left(0, \frac{\lambda_k - (\mu_k + \theta'_k - \theta_k)}{\theta_k}\right) \le m_k \le \frac{\lambda_k}{\theta_k},\\ w_k^{(0)}(m_k) & \textit{if } m_k > \frac{\lambda_k}{\theta_k},\end{cases}\end{aligned} \tag{4.3.1}$$

espresioa du, non

$$\begin{aligned}w_k^{(1)}(m_k) =&\frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k}\frac{\left(C\left(\frac{\lambda_k - (\mu_k + \theta'_k - \theta_k)}{\theta_k}, 1\right) - C(m_k, 1)\right)}{(\lambda_k - (\mu_k + \theta'_k - \theta_k))/\theta_k - m_k},\\ w_k^{(2)}(m_k) =&\frac{(\lambda_k - \theta_k m_k)\frac{\mathrm{d}}{\mathrm{d}m_k}C_k(m_k,1) + (\theta_k m_k + \mu_k + \theta'_k - \theta_k - \lambda_k)\frac{\mathrm{d}}{\mathrm{d}m_k}C_k(m_k,0)}{\theta_k},\\ w_k^{(0)}(m_k) =&\frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k}\frac{\left(C_k(m_k,0) - C_k\left(\frac{\lambda_k}{\theta_k}, 0\right)\right)}{m_k - \lambda_k/\theta_k}.\end{aligned}$$

Frogapena. (4.3.1) ekuazioa ez-beherakorra da, izan ere, $C_k(m_k, 1)$ edozein $m_k \le m'_k$-rako ganbila eta ez-beherakorra da, $\frac{C_k(m'_k,a) - C_k(m_k,a)}{m'_k - m_k}$, funtzioa ez-beherakorra da m_k-n $a = 0, 1$ bada, eta $C_k(m_k, 0) - C_k(m_k, 1)$ ez-beherakorra izateak $\frac{\mathrm{d}_k C_k(m_k,0)}{\mathrm{d}m_k} \ge \frac{\mathrm{d}_k C_k(m_k,1)}{\mathrm{d}m_k}$ inplikatzen du. (4.3.1) Ekuazioa 4.1. Proposiziotik eta 4.6. Lematik ondorioztatzen da. □

Hurrengo proposizioan (4.1.3) problemaren soluzio *bias*-optimoa aurkeztuko da.

4.6 Proposizioa. *Demagun 4.5. Proposizioko baldintza berdinak betetzen direla.* (4.1.3) *problemaren soluzio optimo bat* (3.2.3) *trantsizio tasekin: $s_k(t) = 1$ $W \le w_k(m_k)$ bada eta $s_k(t) = 0$ $W > w_k(m_k)$ bada da, non $w_k(m_k)$ 4.5. Proposizioan definitu den.*

Frogapena. Frogapena 4.1. Hipotesiak ziurtatzean datza. $f_k^a(m_k) = \lambda_k - (\mu_k + \theta'_k - \theta_k)a - \theta_k m_k$ da $a = 0, 1$ bada. $f_k^a(m_k)$-ren diferentziagarritasuna berehalakoa da, $\bar{s}_k(\bar{m}_k)$ monotonoa izatea $\bar{m}_k$ $[m_k^1, m_k^0]$ tartean dagoenean ere bai, eta $f_k^0(m_k)$ ganbila izatea ere bai. $f_k^1(m_k)$ funtzioak

$$\frac{\mathrm{d}^2 f_k^1(m_k)}{\mathrm{d}m_k^2} = 0,$$

betetzen du edozein $m_k \geq 0$-rako, eta beraz ganbila da m_k-n.

Bestalde, $\bar{s}_k(\bar{m}_k) = (\lambda_k - \theta_k m_k)/(\mu_k + \theta'_k - \theta_k)$ da, beraz,

$$\frac{\mathrm{d}^2}{\mathrm{d}m_k^2}\left(\bar{s}_k(\bar{m}_k)\right) = 0,$$

hau da, $\bar{s}_k(\bar{m}_k)$ ganbila da $\bar{m}_k \in [m_k^1, m_k^0]$ tartean.

$f_k^1(m_k) - f_k^1(\bar{m}_k) \geq (\leq) f_k^0(m_k) - f_k^0(\bar{m}_k)$ inekuazioa edozein $m_k \leq (\geq)\bar{m}_k$-rako eta $\bar{m}_k \in [m_k^1, m_k^0]$ denean, $\theta_k(\bar{m}_k - m_k - \bar{m}_k + m_k) \geq (\leq) 0$ espresio bidez adieraz daiteke edozein $m_k \leq (\geq)\bar{m}_k$-rako eta $\bar{m}_k \in [m_k^1, m_k^0]$ denean, zeina beti betetzen den.

Orduan 4.1. eta 4.7. Proposizioetatik, (4.1.3) problemaren soluzio optimo bat $s_k(t) = 1$ $W \leq w_k(m_k)$ bada eta $s_k(t) = 0$ $W > w_k(m_k)$ bada da. □

Ohartu, mantentze-kostuak linealak direnean, indize fluidoa egoerarekiko askea dela eta 3.3. Proposizioan lortu den eredu estokastikoarentzako indizearekin bat datorrela. Demagun orain $C_k(m_k, a_k) = C_k(m_k)$, hau da, mantentze-kostuak sisteman dauden bezeroek bakarrik eragiten dituztela. Kasu horretan indize fluidoa

$$w_k^{(2)}(m_k) = \frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k}\frac{\mathrm{d}}{\mathrm{d}m_k}C_k(m_k),$$

idatz daiteke, zeina $C'(m)\mu/\theta$ indize politikarekin bat datorren $\theta'_k = \theta_k$ kasuan. Erregela hau $c\mu/\theta$-Orokortua izenez deituko da ($Gc\mu/\theta$). $w_k^{(1)}(m_k)$ eta $w_k^{(0)}(m_k)$ gaiak

$$\begin{aligned} w_k^{(1)}(m_k) &= \frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k}\frac{(C_k((\lambda_k - (\mu_k + \theta'_k - \theta_k))/\theta_k) - C_k(m_k))}{(\lambda_k - (\mu_k + \theta'_k - \theta_k))/\theta_k - m_k}, \\ w_k^{(3)}(m_k) &= \frac{(\mu_k + \theta'_k - \theta_k)}{\theta_k}\frac{(C_k(m_k) - C_k(\lambda_k/\theta_k))}{m_k - \lambda_k/\theta_k}, \end{aligned}$$

dira. [26] artikuluan lehen mailako diferentzietan oinarritutako indize politikak proposatu dira eta bertan frogatu da uzterik eta iritsierarik gabeko ilara klase-anitz batean indize politika hauek lortzen dutela kosturik txikiena.

Whittle indizearen eta indize fluidoaren ebaluatze numerikoak uzteak dituen eredurako 3. Kapituluan aurkeztu dira, beraz irakurleak 3. Kapitulura jo dezake politika hauen errendimenduari buruz ikasteko.

4.3.2 Haririk gabeko sare batean *scheduling* oportunista

Sekzio honetan haririk gabeko sare bat kontsideratu da K bezero klasek elkarbanatzen dutena. k klaseko bezeroak λ_k parametroko Poisson prozesu bat jarraituz iristen dira sistemara eta zerbitzuak μ_k parametroko banaketa esponentziala jarraitzen du. Edozein unetan, estazio-baseak informazioa bidali diezaioke sisteman dagoen edozein bezerori. k bezeroaren kanal kalitatea beste bezeroen kalitatearekiko askea da eta G_k $[0, \gamma_k)$-n banaketa uniformeko zorizko algai batekin irudika daiteke. *Scheduling* oportunista dela-eta, k klasea zerbitzatzean edukiera $G_{k,1}, \ldots, G_{k,N_k}$ askeak eta berdinki banatuak dauden zorizko aldagaien maximoa da, ikusi [28]. Beraz, itxarotako edukiera $\mathbb{E}(\max(G_{k,1},\ldots,G_{k,N_k})) = \gamma_k N_k(t)/(N_k(t)+1)$ definitzen du. Irteera tasa gisa $\mu_k(N_k) = \mu_k N_k/(N_k+1)$ har daiteke, non $\mu_k := \tilde{\mu}_k \gamma_k$ den. Markov Erabaki prozesu hau hurrengo trantsizio tasek definitzen dute:

$$b_k^a(n_k) = \lambda_k, \text{ eta } d_k^a(n_k) = \mu_k \frac{n_k}{n_k+1} a, \tag{4.3.2}$$

non $a = 1$ ($a = 0$) akzioaren adierazle den, ikusi 2.1. Irudia. Sistema egonkorra izan dadin $\rho := \sum_{k=1}^{K} \lambda_k/\mu_k < 1$ onartuko da.

Helburua batez besteko mantentze-kostua minimizatzea da, non $C_k(N_k, a)$ N_k k klaseko bezero daudeneko mantentze-kostua den. Ohartu $C_k(N_k, a) = C_k(N_k)$ kostuak sisteman dauden bezeroei egiten diela erreferentzia, eta $C_k(N_k, a) = C_k((N_k - a)^+)$ kostuak ilaran dauden bezeroei.

Testuinguru honetan (2.3.3) problemaren soluzio optimoa 0-1 egiturako atari-politika da. Azken hau 2.1. Proposiziotik ondorioztatzen da.

k klaseko bezeroen oreka egoerako probabilitateak (jaiotza-eta-heriotza prozesuen ohiko formulak erabiliz lortu direnak) n_k politikapean

$$\begin{aligned}
\pi_k^{n_k}(m_k) &= 0, \forall m_k \le n_k - 1,\\
\pi_k^{n_k}(m_k) &= \left(\frac{\lambda_k}{\mu_k}\right)^{m_k-n_k} \frac{m_k+1}{n_k+1} \pi^{n_k}(n_k), \forall m_k \ge n_k+1,\\
\pi_k^{n_k}(n_k) &= 1/\left(1 + \frac{1}{n_k+1} \sum_{i=1}^{\infty} \left(\frac{\lambda_k}{\mu_k}\right)^i (n_k+1+i)\right),
\end{aligned}$$

formulak definitzen ditu. Orain $\sum_{i=0}^{n} \pi_k^n(i)$ funtzioa n-n hertsiki gorakorra dela ikusiko da, hau da, $\pi_k^n(n) \le \pi_k^{n+1}(n+1)$. Hori egin ahal izateko $\pi_k^n(n) \le \pi_k^{n+1}(n+1)$ funtzioa $(\frac{1}{n+1} - \frac{1}{n+2}) \sum_{i=1}^{\infty} (\frac{\lambda_k}{\mu_k})^i, i \ge 0$ bezala idatz daitekela ohartu behar da, edozein n-tarako betetzen dena. 2.2. Proposiziotik (eta beraz 2.1. Korolariotik) problemak indize gaitasuna duela ondorioztatzen da eta Whittle indizea (2.3.6) ekuazioak definitzen duela (2.3.6) ez-beherakorra denean. (2.3.6) ez-beherakorra dela ikusi daiteke numerikoki. Halere, Whittle indize politikaren parametroekiko dependentzia ez da argi geratzen. Hau da indize fluido politika garatzearen motibazio garrantzitsuena, zeinak espresio esplizitu erabilgarri bat duen.

Dinamika fluidoak $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \lambda_k - \mu_k \frac{m_k}{m_k+1} s_k(t)$, dira, non $s_k(t) \in \{0,1\}$ ($s_k(t) = 1$ k estazioa aktibatzen bada). Beraz, $m_k^0 = \infty$ eta $m_k^1 = \lambda_k/(\mu_k - \lambda_k)$, hau da, oreka puntuek $\bar{m}_k \in [m_k^1, \infty)$ betetzen dute. 4.1. Proposiziotik indize fluidoa garatu daiteke, (4.1.3) problemaren *bias*-kostua minimizatzen duen politika deskribatzen duena.

Hurrengo proposizioan indize fluidoa garatu da.

4.7 Proposizioa. *Demagun $C_k(m_k, a)$ diferentziagarria, ganbila eta ez-beherakorra dela m_k-n. Demagun, $C_k(m_k, 0) - C_k(m_k, 1)$ ganbila eta ez-beherakorra dela ere m_k-n. Orduan, indize fluidoa ez-beherakorra da eta*

$$w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + \begin{cases} w_k^{(1)}(m_k) & \textit{if } m_k < \lambda_k/(\mu_k - \lambda_k),\\ w_k^{(2)}(m_k) & \textit{if } \lambda_k/(\mu_k - \lambda_k) \le m_k, \end{cases} \tag{4.3.3}$$

espresioa du, non

$$w_k^{(1)}(m_k) = \mu_k m_k \frac{C_k((\lambda_k/(\mu_k-\lambda_k),1) - C_k(m_k,1)}{\lambda_k - (\mu_k-\lambda_k)m_k},$$
$$w_k^{(2)}(m_k) = m_k(m_k+1)\left(\frac{\mathrm{d}C_k(m_k,1)}{\mathrm{d}m_k} - \frac{\mathrm{d}C_k(m_k,0)}{\mathrm{d}m_k}\right) + \frac{m_k^2\mu_k}{\lambda_k}\frac{\mathrm{d}C_k(m_k,0)}{\mathrm{d}m_k}.$$

Frogapena. (4.3.3) Ekuazioa ez-beherakorra izatea, $C_k(m_k,1)$ ganbila eta ez-beherakorra izatetik edozein $m_k \le m'_k$ denean, $\frac{C_k(m'_k,1)-C_k(m_k,1)}{m'_k-m_k}$, ez-beherakorra izatetik m_k-n eta $C_k(m_k,0)-C_k(m_k,1)$ ganbila eta ez-beherakorra izatetik ondorioztatzen da. Izan ere, azken honek $\frac{\mathrm{d}_kC_k(m_k,0)}{\mathrm{d}m_k} - \frac{\mathrm{d}_kC_k(m_k,1)}{\mathrm{d}m_k} \ge 0$ inplikatzen du eta ez-beherakorra da. (4.3.3) Ekuazioa 4.1. Proposiziotik eta 4.8. Lematik ondorioztatzen da. □

Hurrengo proposizioan (4.1.3) problemaren soluzio *bias*-optimoa aurkeztuko da.

4.8 Proposizioa. *Demagun $C_k(m_k,a)$ diferentziagarria, ganbila eta ez-beherakorra dela m_k-n. Demagun $C_k(m_k,0)-C_k(m_k,1)$ ganbila eta ez-beherakorra dela m_k-n.* (4.1.3) *Problemaren soluzio optimo bat* (4.3.2) *trantsizio tasentzat: $s_k(t)=1$ da $W \le w_k(m_k)$ bada eta $s_k(t)=0$ $W > w_k(m_k)$ bada, non $w_k(m_k)$ 4.7. Proposizioan definitu den.*

Frogapena. Frogapena 4.1. Hipotesiak betetzen direla ziurtatzean datza. $f_k^a(m_k) = \lambda_k - \mu_k\frac{m_k}{m_k+1}a$ da $a=0,1$ bada. $f_k^a(m_k)$ diferentziagarria izatea berehalakoa da, eta $\bar{s}_k(\bar{m}_k)$ monotonoa izatea $\bar{m}_k$ $[m_k^1, m_k^0]$ tartean dagoenean, eta $f_k^0(m_k)$ ganbila izatea ere. $f_k^1(m_k)$ funtzioak

$$\frac{\mathrm{d}^2 f_k^1(m_k)}{\mathrm{d}m_k^2} = \mu_k\frac{2}{(m_k+1)^3} \ge 0,$$

betetzen du, edozein $m_k \ge 0$-tarako, eta beraz ganbila da m_k-n.

$\bar{s}_k(\bar{m}_k) = \lambda_k(\bar{m}_k+1)/\mu_k\bar{m}_k$ denez

$$\frac{\mathrm{d}^2}{\mathrm{d}m_k^2}\left(\bar{s}_k(\bar{m}_k)\right) = 2\frac{\lambda_k}{\mu_k\bar{m}_k^3} \ge 0,$$

betetzen da, hau da $\bar{s}_k(\bar{m}_k)$ ganbila da $\bar{m}_k \in [m_k^1, m_k^0]$ tartean.

$f_k^1(m_k) - f_k^1(\bar{m}_k) \ge (\le) f_k^0(m_k) - f_k^0(\bar{m}_k)$ inekuazioa edozein $m_k \le (\ge)\bar{m}_k$-tarako eta $\bar{m}_k \in [m_k^1, m_k^0]$, $\mu_k(\frac{\bar{m}_k}{\bar{m}_k+1} - \frac{m_k}{m_k+1}) \ge (\le) 0$ espresio gisa idatz daiteke edozein $m_k \le (\ge)\bar{m}_k$-rako eta $\bar{m}_k \in [m_k^1, m_k^0]$, zeina $\frac{m_k}{m_k+1}$ ez-beherakorra delako betetzen den.

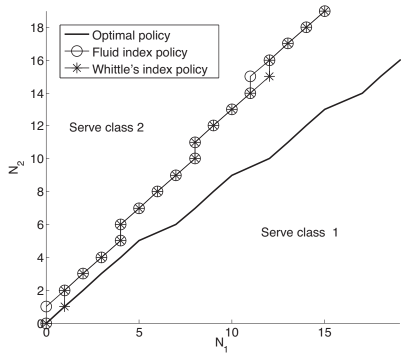
Orduan, 4.1. eta 4.7. Proposizioetatik, (4.1.3). problemaren soluzio optimoa $s_k(t)=1$ da $W \le w_k(m_k)$ bada eta $s_k(t)=0$ $W > w_k(m_k)$ bada. □

Indize fluidoa ez-beherakorra izateak 4.1.3. Sekzioan definitu den indize fluidoak klase bati garrantzi gehiago emango dio berari dagokion ilara hazten den heinean.

(4.3.3) Ekuazioan indize fluidoak duen espresioari esker hainbat ondorio atera daitezke sistemare portaerarekiko. Adibidez, demagun mantentze-kostuak linealak direla $C_k(m,0) = C_k(m,1) = c_k m$, eta $\lambda_k = \lambda\delta_k$. Orduan $w_k(m_k) = c_k m_k/(1-\rho_k)$ $m_k < \lambda_k/(\mu_k-\lambda_k)$ denean eta $w_k(m_k) = c_k m_k^2/\rho_k$ bestela. Beraz, $\lambda \downarrow 0$ den heinean, jatorritik gertu dauden puntuetan $c_k m_k/(1-\rho_k)$ politika erabiliko da eta bestela $c_k\mu_k m_k^2/\delta_k$ politika.

Taula 4.1: 1. Adibidea: Soluzio optimoarekiko errore erlatiboa %-tan.

ρ	0.1	0.2	0.3	0.4
Whittle indize politika	0.20289	1.16215	2.54794	3.54934
Indize fluido politika	0.20289	1.16215	2.55440	3.54936
ρ	0.5	0.6	0.7	0.8
Whittle indize politika	3.52057	2.54793	1.56715	0.66077
Indize fluido politika	3.52098	2.55439	1.60799	0.75140



Irudia 4.2: Trukatze-funtzioak Whittle indizea, indize fluidoa eta soluzio optimoarentzat.

Beheko adibidean bi indize politiken errendimendua aztertuko da numerikoki eta soluzio optimoarekiko konparatuko da.

1. Adibidea Izan bedi bi bezero klase non $\mu_1 = 16, \mu_2 = 27$, eta $\lambda_1/\mu_1 = \rho/2$, $\lambda_2/\mu_2 = \rho/2$. Demagun mantentze-kostuak $C_k(n,a) = c_k(n-a)^2 + b_k(n-a)$ direla $k \in \{1,2\}$, eta $b_1 = 0.1, b_2 = 1, c_1 = 2$ eta $c_2 = 1.5$ hartuz. Hau da, kostuak kuadratikoak dira. Errore erlatiboa kalkulatu da Whittle indize eta indize fluido politikentzat politika optimoarekiko, ikusi 4.1. Taula. Bertan ikusten da bi indize politikak ia optimoak direla hainbat lan kargentzat.

4.2. Irudian politika optimoak hartu dituen erabakiak marraztu dira, Whittle indize politika, eta indize fluido politika $\rho = 0.5$ denean. Hiru politikak irudian agertzen diren trukatze-funtzioek karakterizatzen dituzte. Trukatze-funtzioaren azpian 1 klaseko bezeroak zerbitzatuko dira eta trukatze-funtzioaren gainetik 2 klaseko bezeroak dute lehentasuna.

Indize fluido eta Whittle indize politikei dagozkien trukatze-funtzioak ia baliokideak dira, eta soluzio optimoaren egitura kualitatiboa atzematen dute.

4.3.3 Blokeatzea/bideratzea kontzientzia energetikoz hornitutako zerbitzari-parketan

K estazio ezberdineko zerbitzari-parke bat da aztergai sekzio honetan, non estazio bakoitzak zerbitzari bat duen, ikusi 2.1. Irudia (eskuinaldea). Bezeroak λ tasako Poisson prozesu baten bidez iristen dira sistemara eta sistematik blokeatuko den ala zerbitzari batetara bideratuko den erabaki behar da. Zerbitzarien kapazitateak abiadura-doitze erregela bat jarraitzen dute [99], hau da, zenbat eta bezero gehiago izan ilaran orduan eta azkarragoa da zerbitzua. Energia kontsumoa eta zerbitzariaren kapazitatea orekatzea ahalbideratzen du honek. k zerbitzaria N_k egoeran dagoenean zerbitzariaren kapazitatea $c(N_k) := N_k^\alpha$ da, $0 < \alpha < 1$ delarik. Bezeroaren zerbitzu beharrak k estazioan banaketa esponentziala jarraitzen du μ_k parametroarekin. Beraz, irteera tasa $\mu_k(N_k) = \mu_k N_k^\alpha$ da.

Bezero bat sistematik blokeatzen denean D penalizazio bat jasotzen da, beraz, blokeatze kostuak λD tasa batez eragiten dira. Energia kontsumorako ohiko eredu bat $c(N_k)^{1/\alpha}$ da, beraz, N_k egoeran kontsumitutako energia N_k da. Beraz, kostu gisa $C_k(N_k, a) = C_k(N_k) + \beta_k N_k + D\lambda(1-a)$ definituko da, non $C_k(N_k)$ mantentze-kostua den eta $\beta_k \geq 0$ energiaren kontsumoaren adierazle den. Izan bedi $C_k(N_k)$ ganbila. Helburua blokeatze/bideratze politika optimoa aurkitzea da batez besteko mantentze-kostua, kontsumo energetikoa eta penalizazioa minimizatzeko. Trafikoaren karga banaketarako soluzio optimoak bezero bat ilara handia duen zerbitzari batera bideratzearen eta ilara txikia duen zerbitzari batera bideratzearen arteko oreka aurkitu behar du. Ilara handira bideratzeak energia kontsumo handia suposatzen du, eta ilara txikira bideratzeak zerbitzariaren abiadura txikia izatea inplikatzen du. Hau oso optimizazio problema konplexua da. Tesi honetan garatu diren bi indize politikak errendimendu ona erakusten dutela ikusiko da.

Markov kateak hurrengo trantsizioak ditu:

$$b_k^a(m_k) = \lambda a, \text{ eta } d_k^a(m_k) = \mu_k m_k^\alpha, \tag{4.3.4}$$

non $a = 0$ ($a = 1$) akzioak bezero bat k zerbitzarian blokeatzea (onartzea) adierazten duen.

Eredu honetan, 1-0 egiturako atari-politikak dira (2.3.3) problemaren soluzio optimoa. Frogapena 2.1. Proposiziotik ondorioztatzen da.

Oreka egoerako probabilitateak (jaiotza-eta-heriotza motako prozesuen ohiko formulak erabiliz lortu dena) k klaseko bezeroentzak n_k politikapean

$$\begin{aligned}
\pi_k^{n_k}(m_k) &= 0, \forall m_k \geq n_k + 2, \\
\pi_k^{n_k}(m_k) &= \frac{\lambda^{m_k}}{\mu_k^{m_k} \prod_{i=1}^{m_k} i^\alpha} \pi^{n_k}(0), \forall m_k \leq n_k + 1, \\
\pi_k^{n_k}(0) &= \left(\sum_{m_k=0}^{n_k+1} \frac{\lambda^{m_k}}{\mu_k^{m_k} \prod_{i=1}^{m_k} i^\alpha} \right)^{-1},
\end{aligned}$$

dira. Orain $\sum_{i=0}^n \pi_k^n(i)$ funtzioa n-n hertsiki gorakorra dela ikusiko da, edo baliokideki, $\pi_k^n(n+1)$ hertsiki beherakorra dela n-n. $\pi_k^n(n+1) \leq \pi^{n-1}(n)$ frogatu nahi da, zeina

$$\begin{aligned}
\pi^{n-1}(0) \geq \frac{\lambda}{\mu(n+1)^\alpha} \pi^n(0) &\Leftrightarrow \sum_{m_k=0}^{n+1} \frac{\lambda^{m_k}}{\mu_k^{m_k} \prod_{i=1}^{m_k} i^\alpha} \geq \frac{\lambda}{\mu(n+1)^\alpha} \sum_{m_k=0}^{n} \frac{\lambda^{m_k}}{\mu_k^{m_k} \prod_{i=1}^{m_k} i^\alpha} \\
&\Leftrightarrow \sum_{m_k=1}^{n} \frac{\lambda^{m_k}}{\mu_k^{m_k} \prod_{i=1}^{m_k-1} i^\alpha} \left(\frac{1}{m_k^\alpha} - \frac{1}{(n+1)^\alpha} \right) + 1 \geq 0,
\end{aligned}$$

frogatzearen baliokidea den. Azken inekuazioa $1/m_k^\alpha - 1/(n+1)^\alpha \geq 0$ izateak inplikatzen du. 2.2. Proposiziotik (eta ondorioz 2.1. Korolariotik) problemak indize gaitasuna duela esan daiteke eta Whittle indizea (2.3.6) ekuazioak definitzen duela (2.3.6) ez-beherakorra denean. (2.3.6) Ekuazioa numerikoki kalkulatu daiteke eta ez-beherakorra dela frogatu daiteke. Bertsio fluidoaren soluzio optimoa ere 1-0 egiturako atari-politika da (indize fluidoa beherakorra baita).

Lehenik eta behin indize fluido politika garatuko da eredu honentzat. Dinamika fluidoak $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \lambda s_k(t) - \mu_k m_k^\alpha$, ekuazioak definitzen ditu non $s_k(t) \in \{0,1\}$. $m_k^0 = 0$, da eta $m_k^1 = (\lambda/\mu_k)^{1/\alpha}$, hau da, oreka puntuak hurrengo tartean daude $\bar{m}_k \in [0, m_k^1]$.

Hurrengo proposizioan indize fluidoa garatuko da. Frogapena 4.1. Proposiziotik eta 4.1. Lematik ondorioztatzen da.

4.9 Proposizioa. *Demagun $C_k(m_k)$ polinomio bat dela P mailakoa $P > \alpha$ bada. Orduan, indize fluidoa ez-gorakorra da eta*

$$w_k(m_k) = D\lambda + \begin{cases} w_k^{(2)}(m_k) & \text{if } 0 \leq m_k \leq (\lambda/\mu_k)^{\alpha^{-1}}, \\ w_k^{(1)}(m_k) & \text{if } (\lambda/\mu_k)^{\alpha^{-1}} < m_k, \end{cases}$$

da, non

$$w_k^{(2)}(m_k) = -\frac{\lambda\alpha^{-1} m_k}{\mu_k m_k^\alpha} \frac{\mathrm{d}\tilde{C}_k(m_k)}{\mathrm{d}m_k}$$
$$w_k^{(1)}(m_k) = -\lambda \frac{(\tilde{C}_k((\lambda/\mu_k)^{\alpha^{-1}}) - \tilde{C}_k(m_k))}{\lambda - \mu_k \min(T, m_k^\alpha)},$$

eta $\tilde{C}_k(m_k) = C_k(m_k) + \beta_k m_k$.

Hurrengo proposizioan (4.1.3) problemarentzat soluzio optimoa aurkeztuko da.

4.10 Proposizioa. *Demagun $C_k(m_k)$ polinomio bat dela P mailakoa non $P > \alpha$.* (4.1.3) *problemaren soluzio optimo bat* (4.3.4) *trantsizioekin hurrengoa da: $s_k(t) = 1$ $W \leq w_k(m_k)$ bada eta $s_k(t) = 0$ $W > w_k(m_k)$ bada, non $w_k(m_k)$ 4.9. Proposizioan definitu den.*

Frogapena. Frogapena 4.1. Hipotesiak ziurtatzean datza. $f_k^a(m_k) = \lambda a - \mu_k m_k^\alpha$ da edozein $a \in \{0,1\}$-rako eta $\alpha < 1$ da. $f_k^a(m_k)$ diferentziagarria izatea $\bar{s}_k(\bar{m}_k)$ $[m_k^0, m_k^1]$ tartean monotonoa izatetik ondorioztatzen da. Bestalde, $f_k^a(m_k)$ funtzioak

$$\frac{\mathrm{d}^2 f_k^a(m_k)}{\mathrm{d}m_k^2} = -\alpha(\alpha-1)\mu_k m_k^{\alpha-2} \geq 0,$$

betetzen du edozein $m_k \geq 0$ denean, $\alpha < 1$ baita. Beraz, $f_k^a(m_k)$ ganbila da m_k-n $a = 0, 1$ bada.

$1 - \bar{s}_k(\bar{m}_k) = 1 - \mu_k m_k^\alpha/\lambda$ denez

$$\frac{\mathrm{d}^2}{\mathrm{d}m_k^2}(1 - \bar{s}_k(\bar{m}_k)) = -\frac{\mu_k\alpha(\alpha-1)\bar{m}_k^{\alpha-2}}{\lambda} \geq 0,$$

edozein $\bar{m}_k \in [m_k^0, m_k^1]$-rako, $\alpha < 1$ delako. Beraz, $1 - \bar{s}_k(\bar{m}_k)$ ganbila da $\bar{m}_k$-n.

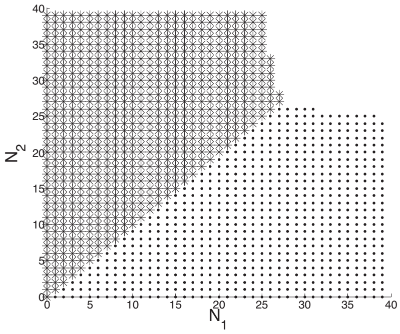
$f_k^1(m_k) - f_k^1(\bar{m}_k) \leq (\geq) f_k^0(m_k) - f_k^0(\bar{m}_k)$ inekuazioa $\mu_k(\bar{m}_k^\alpha - m_k^\alpha) \leq (\geq)\mu_k(\bar{m}_k^\alpha - m_k^\alpha)$ bezala idatz daiteke, edozein m_k-rako betetzen dena eta edozein $\bar{m}_k$-rako.

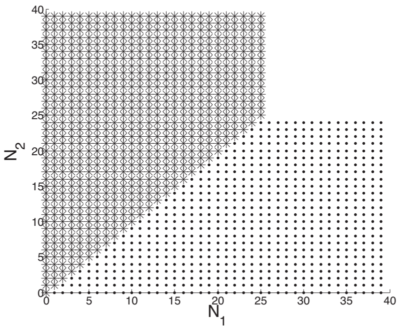
Orduan 4.1. eta 4.9. Proposizioetatik (4.1.3) problemaren soluzio optimoa $s_k(t) = 1$ dela $W \leq w_k(m_k)$ denean eta $s_k(t) = 0$ dela $W > w_k(m_k)$ denean ondorioztatzen da. □

Indize fluidoa ez-gorakorra izateak indize fluido politikak ilara txikiak dituzten zerbitzarietara bideratuko dituela bezero berriak inplikatzen du. Indize fluido politikak bezeroak indize positiboak dituzten

Taula 4.2: 2. Adibidea: Errore erlatiboa %-tan.

ρ	0.1	0.3	0.5
Indize fluidoa	0.08704×10^{-7}	0.16036×10^{-7}	0.13968×10^{-7}
Whittle indizea	0.08704×10^{-7}	0.16036×10^{-7}	0.13968×10^{-7}
ρ	0.7	0.9	1.1
Indize fluidoa	0.06279×10^{-7}	0.08210×10^{-7}	0.06124×10^{-7}
Whittle indizea	0.06279×10^{-7}	0.08210×10^{-7}	0.06124×10^{-7}
ρ	1.5	2	2.5
Indize fluidoa	0.01872×10^{-7}	0.06099×10^{-7}	0.10921×10^{-7}
Whittle indizea	0.01872×10^{-7}	0.06099×10^{-7}	0.07110×10^{-7}





Irudia 4.3: “*” (“.”)-ko eremuan 1 klaseko (2 klaseko) bezeroak dute lehentasuna eta ikurrik gabeko eremuan bezeroak blokeatu egiten dira. Ezkerraldean: politika optimoa. Eskuinaldean: Whitte-en indize politika eta indize fluido politikak.

zerbitzarietara bakarrik bidaliko ditu, beraz, existitzen da $\overline{N}_k$ zeinarentzat $N_k \geq \overline{N}_k$ eta bezeroak blokeatuak izango diren.

Aurreko sekzioan bezela 4.9. Proposizioa erabili da kasu partikularretan ondorioak ateratzeko. Adibidez, demagun matentze kostuak linealak direla eta $C_k(m) = c_k m$. Orduan, $\lambda \uparrow \infty$, den heinean $w_k(m_k)$ indizeak $D\lambda + w_k^{(2)}(m_k)$ espresioa hartzen du non $w_k^{(2)}(m_k) = -\lambda c_k \frac{m_k^{1-\alpha}}{\mu_k \alpha}$, beraz lehentasuna $c_k \frac{m_k^{1-\alpha}}{\mu_k \alpha}$ indizearen arabera ezarriko da.

Orain adibide bat aurkeztuko da indize politiken errendimendua erakusteko.

2. Adibidea Adibide honetan 2 bezero klase daudela suposatuko da eta iritsiera tasa $\lambda = 18$ dela. Abiadura-doitze tasa $\alpha = 1/2$ dela onartuko da. Demagun kostu funtzioa $C_k(m_k, a) = C_k(m_k) + \beta_k m_k + D\lambda a$ dela, eta $C_k(m_k) = c_k m_k^2$ dela non $c_1 = c_2 = 2$ eta $\beta_1 = 3, \beta_2 = 5$. Bestalde, demagun $D = 25$ dela eta μ_1, μ_2 parametroek $\mu_1 = \mu_2 = 2\lambda/\rho$ betetzen dutela. $M = 1$ finkatuko da, hau da, bezero bat gehienez zerbitzari batetara bidal daiteke. 4.2. Taulan ikusten da Whittle indizea eta indize fluido politiken errendimendua ia optimoa dela hainbat ρ lan kargatarako. Ez hori bakarrik, 4.3. Irudian politika optimoa eta Whittle indize politika irudikatu dira $\rho = 2.5$ denan. Indize fluido politikak kasu honetan Whittle indize politikarekin bat egiten du eta soluzio optimoaren egitura kualitatiboa jasotzen du.

4.3.4 Inbentarioen kudeaketa itemak galkorrak direnean

Eredu honetan K item ezberdin ekoiz ditzakeen makina bat kontsideratuko da. Problema hau 2.1. Irudian (eskuinaldean) ilustratu diren ereduen adibide bat da. k klaseko itemaren eskaerak Poisson prozesu bat jarraitzen du λ_k tasakoa. Makinak item bakarra ekoiz dezake une bakoitzean, eta ekoizpen denborak banaketa esponentziala jarraitzen du μ_k parametroarekin. Itemak saldu bitartean gordeta gelditzen dira, baina galkorrak direnez, saldu aurretik usteldu edo gal daitezke. Galtze-prozesuak banaketa esponentziala jasaten du θ_k parametroarekin. Makinak item bat ekoiztu ala ez erabaki behar du. Makinak item bat ekoiztea erabakitzen badu zein klaseko itema ekoiztu erabaki behar du. Problema hau itemak galkorrak ez diren kasuan [90] artikuluan aztertu da. $N_k(t)$ aldagaiak inbentarioan dauden k klaseko item kopurua adieraziko du. MEP-a hurrengo trantsizio tasek definitzen dute:

$$b_k^a(m_k) = \mu_k a \text{ eta } d_k^a(m_k) = \lambda_k + \theta_k m_k, \tag{4.3.5}$$

non $a = 1(a = 0)$ akzioak item bat ekoiztea (ez-ekoiztea) adierazten duen. k klaseko item batentzat eskaera bat iristen denean eta $N_k(t) = 0$ bada, *i.e.*, ez badago k klaseko itemik gordeta, orduan salmenta bat galduko da. Azken honek galtzen den salmenta bakoitzeko kostu bat eragiten du $D_k > 0$. k klaseko item bat galtzen denean (epea amaitzen denean), item bakoitzeko kostu bat ordaindu behar da δ_k. Sistemak $c_k(m, a)$ denbora unitateko kostu bat eragiten du gordeta dauden k klaseko m itemengatik, $a = 0, 1$ bada. Orduan, denbora unitateko eragiten den kostua k klaseko $C_k(N_k(t), a) = c_k(N_k(t), a) + \delta_k\theta_k N_k(t) + \lambda_k D_k \mathbb{1}_{\{N_k(t)=0\}}$ da. Helburua eredu erlaxatu estokastikoan batez besteko kostua minimizatzea da, *i.e.*,

$$\mathbb{E}(\tilde{C}_k(N_k^\phi, S_k^\phi(N_k^\phi))) + D_k\lambda_k\pi_k^\phi(0),$$

non $\tilde{C}_k(m, a) = c_k(m, a) + \delta_k\theta_k m$, eta N_k^ϕ oreka egoeran dauden k klaseko itemak diren ϕ politikapean.

Kontextu honetan (2.3.3) problemaren soluzio optimoa 1-0 egitureko atari-politika bat da. Frogapena 2.1. Proposiziotik ondorioztatzen da.

Oreka egoerako probabilitateak $\pi_k^{n_k}(m_k)$ n_k atari-politikapean

$$\pi_k^{n_k}(m_k) = \frac{\mu_k^{m_k}}{\prod_{i=1}^{m_k}(\lambda_k + \theta i)}\pi^{n_k}(0), \text{ edozein } m_k \le n_k + 1\text{-rako},$$
$$\pi_k^{n_k}(m_k) = 0, \text{ edozein } m_k \ge n_k + 2\text{-rako},$$

dira, non $\pi_k^{n_k}(0) = \left(\sum_{m=0}^{n_k+1} \frac{\mu_k^m}{\prod_{i=1}^m(\lambda_k+\theta_k i)}\right)^{-1}$. Orain $\pi^n(n+1)$ hertsiki beherakorra dela ikusiko da n-n 2.2. Proposizioan indize gaitasunerako eskatzen baita, Beraz, $\pi^n(n+1) \le \pi^{n-1}(n)$ frogatu behar da edozein $n \ge 0$-rako, hau da,

$$\frac{\mu_k^{n+1}}{\prod_{i=1}^{n+1}(\lambda_k + \theta_k i)}\pi_k^n(0) \le \frac{\mu_k^n}{\prod_{i=1}^{n}(\lambda_k + \theta_k i)}\pi_k^{n-1}(0)$$
$$\Leftrightarrow \frac{\mu_k}{\lambda_k + \theta_k(n+1)}\sum_{m=0}^{n}\frac{\mu_k^m}{\prod_{i=1}^{m}(\lambda_k + \theta_k i)} \le \sum_{m=0}^{n+1}\frac{\mu_k^m}{\prod_{i=1}^{m}(\lambda_k + \theta_k i)}$$
$$\Leftrightarrow \sum_{m=1}^{n}\frac{\mu_k^m}{\prod_{i=1}^{m-1}(\lambda_k + \theta_k i)}\left(\frac{1}{\lambda_k + \theta_k m} - \frac{1}{\lambda_k + \theta_k(n+1)}\right) + 1 \ge 0.$$

Azken inekuazio hau $(\frac{1}{\lambda_k+\theta_k m} - \frac{1}{\lambda_k+\theta_k(n+1)}) \geq 0$ delako betetzen da. Beraz, 2.2. Proposiziotik Whittle indizea (2.3.6) ekuazioak definitzen duela ondorioztatzen da (2.3.6) ez-gorakorra bada. (2.3.6) Ekuazioa numerikoki kalkula daiteke eta ez-beherakorra dela ikusi daiteke, Halere, ez da erraza parametroekiko Whittle indize politikak duen influentzia. Hori dela eta indize fluido politika garatu da, zeinak espresio esplizitua duen.

Dinamika fluidoak $\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \mu_k s_k(t) - \lambda_k - \theta_k m_k$ espresioak ematen ditu edozein $m_k \geq 0$-rako. Demagun $\lambda_k < \mu_k$ eta beraz 4.1.2. Sekzioan egindako hipotesietatik, *i.e.* $m_k^a = 0$ $f_k^a(m_k) < 0$ bada eta $m_k \geq 0$, orduan $m_k^0 = \max\{0, -\lambda_k/\theta_k\} = 0$ eta $m_k^1 = \max\{\frac{\mu_k - \lambda_k}{\theta_k}, 0\} = \frac{\mu_k-\lambda_k}{\theta_k}$. Sistema estokastiko honetan ez zerbitzatzea erabakitzen den unetik ez da hortik gorako egoerarik bisitatuko eta beraz n_k politikapean, salmenta galerei dagokien batez besteko kostua $D_k \lambda_k \pi_k^{n_k}(0)$ da. Hau da batez bestean ordaindu den kostua, eta beraz, eredu fluidoaren funtzio objektiboak ez du atzematen, sistema fluidoan atari-politikapena ez baita inoiz 0 egoera ukitzen. Ordez, hurrengo kostua onartuko da, $C_k(m_k, a) := c_k(m_k, a) + \theta_k \delta_k m_k + \lambda_k D_k(1-a)$.

Indize fluidoa hurrengo proposizioan garatuko da.

4.11 Proposizioa. *Demagun $C_k(m_k, a)$ ganbila, diferentziagarria eta ez-beherakorra dela, eta $C_k(m_k, 1) - C_k(m_k, 0)$ ganbila eta ez-beherakorra. Orduan, indize fluidoa ez-gorakorra da eta*

$$w_k(m_k) = C_k(m_k, 0) - C_k(m_k, 1) + \begin{cases} w_k^{(2)}(m_k), & \textit{if } m_k \leq (\mu_k - \lambda_k)/\theta_k \\ w_k^{(1)}(m_k), & \textit{if } m_k > (\mu_k - \lambda_k)/\theta_k, \end{cases} \tag{4.3.6}$$

espresioa du, non

$$w_k^{(2)}(m_k) = \left(-m_k - \frac{\lambda_k}{\theta_k}\right)\left(\frac{\mathrm{d}C_k(m_k,1)}{\mathrm{d}m_k} - \frac{\mathrm{d}C_k(m_k,0)}{\mathrm{d}m_k}\right) - \frac{\mu_k}{\theta_k}\frac{\mathrm{d}C_k(m_k,0)}{\mathrm{d}m_k},$$
$$w_k^{(1)}(m_k) = -\frac{\mu_k}{\theta_k}\left(\frac{C_k(m_k,1) - C_k((\mu_k-\lambda_k)/\theta_k, 1)}{m_k - (\mu_k-\lambda_k)/\theta_k}\right).$$

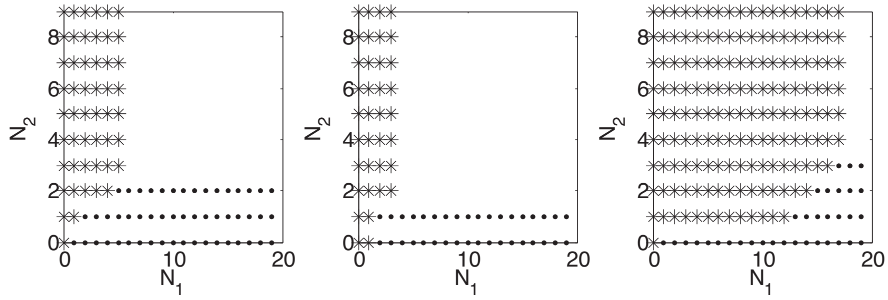
Frogapena. (4.3.6) Ekuazioa 4.1. Proposiziotik eta 4.12. Lematik ondorioztatzen da. Indizea ez-beherakorra izateak $C_k(m_k, 1)$ ganbila eta ez-beherakorra, $\frac{C_k(m_k', 1) - C_k(m_k, 1)}{m_k' - m_k}$, ez beherakorra, eta $C_k(m_k, 1) - C_k(m_k, 0)$ ganbila eta ez-beherakorra izateak inplikatzen dute. Izan ere, $\frac{\mathrm{d}_k C_k(m_k,1)}{\mathrm{d}m_k} - \frac{\mathrm{d}_k C_k(m_k,0)}{\mathrm{d}m_k} \geq 0$ ez-beherakorra izatea inplikatzen dute. □

Orain (4.1.3) problemaren soluzio optimoa aurkeztuko da.

4.12 Proposizioa. *Demagun $C_k(m_k, a)$ ganbila, diferentziagarria eta ez-beherakorra dela, eta $C_k(m_k, 1) - C_k(m_k, 0)$ ganbila eta ez-beherakorra.* (4.1.3) *problemaren soluzio optimoa trantsizio tasak* (4.3.5) *ekuazioak definitzen dituenean: $s_k(t) = 1$ da $W \leq w_k(m_k)$ bada eta $s_k(t) = 0$ $W > w_k(m_k)$ bada, non $w_k(m_k)$ 4.11. Proposizioan definitu den.*

Frogapena. Frogapena 4.1. Hipotesiak betetzen direla ziurtatzean datza. $f_k^a(m_k) = \mu_k a - \lambda_k - \theta_k m_k$ da $a = 0, 1$ bada. $f_k^a(m_k)$ diferentziagarria izatea berehalakoa da eta $\bar{s}_k(\bar{m}_k)$ monotonoa izatea ere $[m_k^0, m_k^1]$ tartean. $f_k^a(m_k)$ funtzioak

$$\frac{\mathrm{d}^2 f_k^a(m_k)}{\mathrm{d}m_k^2} = 0,$$



Irudia 4.4: "*" (".")-ko eremuan 1 klaseak (2 klaseak) du lehentasuna eta ikurrik gabeko eremuan bezeroak blokeatu egiten dira. Ezkerraldean: politika optimoa. Erdian: Whittle indize eta indize fluido politikak.

betetzen du edozein $m_k \geq 0$-rako. Beraz, $f_k^a(m_k)$ ganbila da m_k-n $a = 0, 1$ bada.

$1 - \bar{s}_k(\bar{m}_k) = 1 - (\lambda_k + \theta_k m_k)/\mu_k$ da, beraz

$$\frac{\mathrm{d}^2}{\mathrm{d}m_k^2}\left(1 - \bar{s}_k(\bar{m}_k)\right) = 0,$$

edozein $\bar{m}_k \in [m_k^0, m_k^1]$ tartean. Orduan, $1 - \bar{s}_k(\bar{m}_k)$ ganbila da $\bar{m}_k$-n.

$f_k^1(m_k) - f_k^1(\bar{m}_k) \leq (\geq) f_k^0(m_k) - f_k^0(\bar{m}_k)$ inekuazioa $\bar{m}_k - m_k \leq (\geq) \bar{m}_k - m_k$ ere idatz daiteke, zeina edozein m_k eta $\bar{m}_k$-rako betetzen den.

Orduan 4.1. eta 4.9. Proposizioetatik (4.1.3) problemaren soluzio optimoa $s_k(t) = 1$ da $W \leq w_k(m_k)$ eta $s_k(t) = 0$ $W > w_k(m_k)$ bada. □

Indize fluidoa ez-beherakorra izateak k klaseko zenbat eta item gehiago izan orduan eta k klaseko item gutxiako ekoiztuko direla esan nahi du. Indize fluidoa negatiboa bada klase guztientzat orduan indize fluido politikak itemik ez ekoiztea erabakitzen du.

Indize politiken errendimendua aztertuko da hurrengo adibidean.

3. Adibidea Adibide honetan bi item klase daudela suposatu da eta $\mu_1 = 4$ eta $\mu_2 = 5$ direla, makinak itemak ekoiztea erabakitzen badu. Kostu funtzioa $C_k(m_k, a) = C_k(m_k) + \delta_k\theta_k m_k + \lambda_k D_k \pi^{m_k}(0)$ da, eta demagun $C_k(m_k) = c_k m_k^2 + b_k m_k$ non $c_1 = 1, c_2 = 2$ eta $b_1 = b_2 = 1$, eta $\theta_1 = 2, \theta_2 = 2.5$. Bestalde, galkorrak diren itemen kostua $\delta_1 = 0.5, \delta_2 = 3$ da eta salmentak galtzearen kostua $D_1 = 20$, $D_2 = 14$. Itemen eskaera $\lambda_1 = 3.5, \lambda_2 = 4.8$ tasekin iristen dira. Finkatu $M = 1$, hau da, makina batek ekoizten ditu item guztiak. 4.4. Irudian politika optimoa illustratu da Whittle indize politika eta indize fluido politikekin batera $\rho = 2.5$ denean. Whittle indize politika eta indize fluido politikak soluzio optimoaren egitura kualitatiboa atzematen dute. 4.3. Taulan lan karga ezberdinentzat, errore erlatiboa bi heuristikentzat oso txikia da.

Taula 4.3: 3. Adibidea: Errore erlatiboa %-tan.

ρ	0.1	1	1.5	2
Whittle indize politika	1.204×10^{-4}	0.0036	0.0099	0.0285
Indize fluido politika	0.0532	0.9766	2.8506	4.6276
ρ	2.5	3	3.5	4
Whittle indize politika	0.0676	0.1321	0.2242	4.0901
Indize fluido politika	2.5092	0.0814	0.0987	4.1111

4.4 Eranskina

4.4.1 4.1. Proposizioaren frogapena

k-rekiko dependentzia alde batera utziko da frogapen honetan. Orokortasunik galdu gabe $m^0 > m^1$ dela onartu. Notazioa erraztearren onartu $f^1(m^1) = 0$. $f^1(m^1) < 0$ kasurako frogapena antzera egin daiteke, halere, kasu horretan 2 eta 3 kasuak bakarrik aztertu behar dira (ikusi kasuak behean).

Lehenik eta behin 4.1. Proposizioa frogatzeko jarraitu diren pausuak aurkeztuko dira:

- **1. Pausua:** $w(m)$ jarraia dela frogatuko da m-n.
- **2. Pausua:** $EC(\bar{s}, W)$ $\bar{s}$-rekiko ganbila dela ikusiko da. Honek oreka puntu optimoa (m^*, s^*) karakterizatzea ahalbideratzen du, W finko baterako.
- **3. Pausua:** $w(m)$-ri dagokion politikak Hamilton-Jacobi-Bellman ekuazioa [22] betetzen du, W-ren karakterizazioa erabiliz (2. pausuan lortuko dena) eta $w(m)$ jarraia dela erabiliz (1. pausuan) frogatuko dena.

1. Pausua: Lehenik eta behin $w(m)$ funtzio jarraia dela frogatuko da. Horretarako $\lim_{m\uparrow m^1} w^{(1)}(m) = \lim_{m\downarrow m^1} w^{(2)}(m)$ dela frogatuko da. Ohartu

$$\begin{aligned}\lim_{m\uparrow m^1} w^{(1)}(m) =&(f^0(m^1) - f^1(m^1)) \lim_{m\uparrow m^{1-}} \frac{\big(C(m^1,1) - C(m,1)\big)}{f^1(m)} = \left(\frac{\mathrm{d}f^1(m)}{\mathrm{d}m}\bigg|_{m=m^1} \right)^{-1} \\ &\cdot (f^1(m^1) - f^0(m^1)) \left(\frac{\mathrm{d}C(m,1)}{\mathrm{d}m}\bigg|_{m=m^1} \right),\end{aligned}$$

non azken inekuazioan l'Hopital-en erregela erabili den. $f^1(m^1) = 0$ denez

$$\lim_{m\uparrow m^1} w^{(1)}(m) = -f^0(m^1) \left(\frac{\mathrm{d}C(m,1)}{\mathrm{d}m}\bigg|_{m=m^1} \right) \cdot \left(\frac{\mathrm{d}f^1(m)}{\mathrm{d}m}\bigg|_{m=m^1} \right)^{-1},$$

lortzen da. Bestalde,

$$\lim_{m\downarrow m^1} w^{(2)}(m) = -f^0(m^1) \left(\frac{\mathrm{d}C(m^1,1)}{\mathrm{d}m^1} \right) \cdot \left(\frac{\mathrm{d}f^1(m^1)}{\mathrm{d}m^1} \right)^{-1},$$

beraz bi limiteak bat datoz. Antzeko argumentuak erabiliz $\lim_{m\uparrow m^0} w^{(2)}(m) = \lim_{m\downarrow m^0} w^{(0)}(m)$ froga daiteke, zeinak frogapena amaitzen duen.

2. Pausua: W finko baterako $(\bar{m}, \bar{s})$ oreka puntu optimoa karakterizatuko da. Gogoratu $m^0 > m^1$ eta $f^1(m^1) = 0$. Beraz, $(\bar{m}, \bar{s})$ oreka puntu bat $\frac{\mathrm{d}m(t)}{\mathrm{d}t}$-rentzat

$$(1-\bar{s})f^0(\bar{m}) + \bar{s}f^1(\bar{m}) = 0, \tag{4.4.1}$$

betetzen duen puntu bat da non $\bar{s} \in [0,1]$, hau da, orekan $\bar{s}$ denbora frakzio batean akzio aktiboa aukeratu da eta $1-\bar{s}$ denbora frakzio batean akzio pasiboa. Gogoratu

$$EC(\bar{s}, W) := (1-\bar{s})C(\bar{m}, 0) + \bar{s}C(\bar{m}, 1) - (1-\bar{s})W.$$

Funtzio hau ganbila dela frogatuko da $\bar{s}$-n, hau da $\frac{\mathrm{d}^2 EC(\bar{s},W)}{\mathrm{d}\bar{s}^2} \geq 0$. Hainbat kalkulu eta gero

$$\begin{aligned}\frac{\mathrm{d}^2 EC(\bar{s}, W)}{\mathrm{d}\bar{s}^2} =& \left(\frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}} - \frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}}\right) 2\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}} + (1-\bar{s})\left(\frac{\mathrm{d}^2C(\bar{m},0)}{\mathrm{d}\bar{m}^2}\left(\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\right)^2 + \frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}}\frac{\mathrm{d}^2\bar{m}}{\mathrm{d}\bar{s}^2}\right)\\ &+ \bar{s}\left(\frac{\mathrm{d}^2C(\bar{m},1)}{\mathrm{d}\bar{m}^2}\left(\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\right)^2 + \frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}}\frac{\mathrm{d}^2\bar{m}}{\mathrm{d}\bar{s}^2}\right),\end{aligned} \tag{4.4.2}$$

betetzen dela ikus daiteke. (4.4.1) Ekuaziotik $\bar{s} = f^0(\bar{m})/(f^0(\bar{m}) - f^1(\bar{m}))$ lortzen da eta beraz,

$$\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}} = \left(\frac{\mathrm{d}\bar{s}}{\mathrm{d}\bar{m}}\right)^{-1} = \left(\frac{f^0(\bar{m})\frac{\mathrm{d}f^1(\bar{m})}{\mathrm{d}\bar{m}} - f^1(\bar{m})\frac{\mathrm{d}f^0(\bar{m})}{\mathrm{d}\bar{m}}}{(f^0(\bar{m}) - f^1(\bar{m}))^2}\right)^{-1} \leq 0, \tag{4.4.3}$$

edozein $\bar{s} \in [0,1]$. Inekuazioa f^0 eta f^1 funtzio ez-gorakorrak izatetik eta $f^1(\bar{m}) \leq 0$ eta $f^0(\bar{m}) \geq 0$ edozein $\bar{m} \in [m^1, m^0]$ tartean direnez ondorioztatzen da. $\mathrm{d}\bar{m}/\mathrm{d}\bar{s} \leq 0$ eta $\mathrm{d}^2\bar{m}/\mathrm{d}\bar{s}^2 = -\mathrm{d}^2\bar{s}/\mathrm{d}\bar{m}^2\,(\mathrm{d}\bar{m}/\mathrm{d}\bar{s})^3$ direnez, Katearen erregela erabiliz lor daitezkeenak, $\mathrm{d}^2\bar{s}/\mathrm{d}\bar{m}^2 \geq 0 \Leftrightarrow \mathrm{d}^2\bar{m}/\mathrm{d}\bar{s}^2 \geq 0$ lortzen da. Hipotesiz $\bar{s} = f^0(\bar{m})/(f^0(\bar{m}) - f^1(\bar{m}))$ ganbila da $\bar{m}$-n, beraz, $\mathrm{d}^2\bar{m}/\mathrm{d}\bar{s}^2 \geq 0$. $\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}} \leq 0$ eta $\frac{\mathrm{d}^2\bar{m}^2}{\mathrm{d}\bar{s}^2} \geq 0$ direnez, eta (1) $C(m,a)$, $a = 0,1$, ganbila eta ez-beherakorra denez m-n eta (2) $\frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}} \leq \frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}}$ denez (4.4.2) ekuazioa 0 baino handiago edo berdina dela lortzen da eta beraz, $EC(\bar{s}, W)$ ganbila da $\bar{s} \in [0,1]$ tartean.

$EC(\bar{s}, W)$ ganbila izateak $\bar{s} \in [0,1]$ tartean, oreka puntua hiru hauetatik bat izatea inplikatzen du:

- 1. Kasua: $\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \leq 0$ edozein $\bar{s} \in [0,1]$ tartean, beraz, oreka optimoak $s^* = 1, m^* = m^1$ betetzen du.
- 2. Kasua: $\frac{\mathrm{d}EC(s^*,W)}{\mathrm{d}s^*} = 0$, beraz, oreka optimoak $s^* \in [0,1]$ eta $m^* \in [m^1, m^0]$ betetzen du.
- 3. Kasua: $\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \geq 0$ edozein $\bar{s} \in [0,1]$, beraz, oreka puntu optimoak $s^* = 0$, $m^* = m^0$ betetzen du.

Orain ohartu $\frac{\mathrm{d}EC(\bar{s},W)}{\mathrm{d}\bar{s}} \geq 0$, baldin eta soilik baldin

$$W \geq C(\bar{m},0) + C(\bar{m},1) + (1-\bar{s})\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}} + \bar{s}\frac{\mathrm{d}\bar{m}}{\mathrm{d}\bar{s}}\frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}}, \tag{4.4.4}$$

bada, eta $(1-\bar{s})f^0(\bar{m}) + \bar{s}f^1(\bar{m}) = 0$ ordezkatu ondoren eta $\mathrm{d}\bar{m}/\mathrm{d}\bar{s}$-ren (4.4.3) ekuazioan emandako espresioa ordezkatuz, (4.4.4) ekuazioa

$$W \geq C(\bar{m},0) - C(\bar{m},1) + \frac{(f^1(\bar{m}) - f^0(\bar{m}))(f^0(\bar{m})\frac{\mathrm{d}C(\bar{m},1)}{\mathrm{d}\bar{m}} - f^1(\bar{m})\frac{\mathrm{d}C(\bar{m},0)}{\mathrm{d}\bar{m}})}{f^0(\bar{m})\frac{\mathrm{d}f^1(\bar{m})}{\mathrm{d}\bar{m}} - f^1(\bar{m})\frac{\mathrm{d}f^0(\bar{m})}{\mathrm{d}\bar{m}}},$$

ekuazioaren baliokidea dela lortzen da, hau da,

$$W \geq C(\bar{m}, 0) - C(\bar{m}, 1) + w^{(2)}(\bar{m}).$$

Beraz, 3. Kasuan W $W \geq C(\bar{m}, 0) - C(\bar{m}, 1) + w^{(2)}(\bar{m})$ da edozein $\bar{m} \in [m^1, m^0]$, tartean eta bereziki $W \geq w(m^0)$.

Antzeko moduan, 1. Kasuan egoteak $W \leq w(m^1)$ inplikatzen du.

2. Kasuan, $\mathrm{d}EC(s^*, W)/\mathrm{d}s^* = 0$ berdinketatik, $W = C(m^*, 0) - C(m^*, 1) + w^{(2)}(m^*) = w(m^*)$ lortzen da, $m^* \in [m^1, m^0]$ tartean.

3. Pausua: $s(t) \in \{0, 1\}$ denez dozein t denboran, HJB ekuazioa betetzea da ibilbidea optimoa izateko baldintza nahikoa. (4.1.4) berridatziz hurrengo baldintza lortzen da edozein m-rako:

$$0 = \min\{\mathcal{J}_0(m, W), \mathcal{J}_1(m, W)\}, \tag{4.4.5}$$

non

$$\mathcal{J}_0(m, W) = C(m, 0) - W - EC^*(W) + f^0(m)\frac{\partial J(m, W)}{\partial m}, \tag{4.4.6}$$

$$\mathcal{J}_1(m, W) = C(m, 1) - EC^*(W) + f^1(m)\frac{\partial J(m, W)}{\partial m}, \tag{4.4.7}$$

eta $J(m, W)$ funtzioa jarraiak eta diferentziagarriak diren.

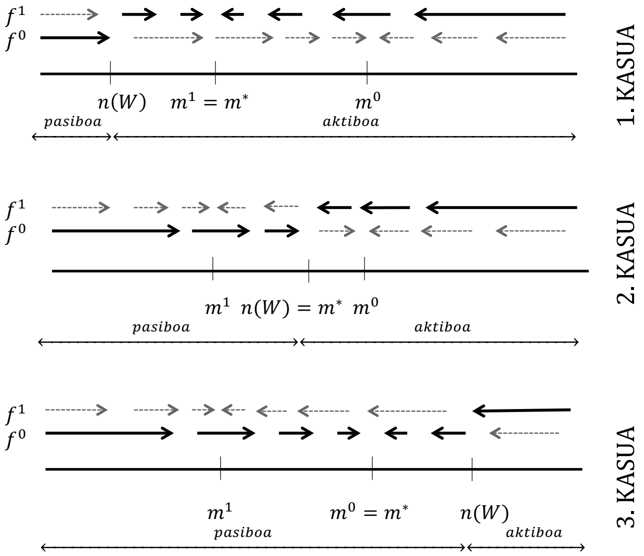
W finko baterako, kontsideratu $s(t) = 0$ den politika edozein m egoeratan zeinentzat $W \geq w(m)$ den eta aktiboa, $s(t) = 1$, edozein m egoeratan zeinentzat $W < w(m)$. Ohartu $w(m)$ ez-beherakorra izateak, hau atari-politika izatea inplikatzen duela. Hau da, existizen da $n(W) \in \mathbb{Z}_+$ zeinentzat $W > w(m)$ edozein $m \leq n(W)$ eta $W \leq w(m)$ edozein $m \geq n(W)$, ikusi 4.5. Irudia. Politika honi $n(W)$ atari-politika deituko zaio. $n(W)$ politikak HJB (4.4.5) ekuazioa betetzen duela ikusi nahi da. Horretarako definitu $J^{n(W)}(m, W)$ W finko batentzat $n(W)$ politikapeko kostu bezela, m hasiera puntu batetik oreka punturarte, hau da,

$$\begin{aligned} J^{n(W)}(m, W) = &\int_0^{t_0} \Big(C(m^{n(W)}(t), s_0) - W(1 - s_0) - EC^*(W)\Big)\mathrm{d}t \\ &+ \int_{t_0}^{\infty} C(m^{n(W)}(t), s_1) - W(1 - s_1) - EC^*(W)\mathrm{d}t, \end{aligned} \tag{4.4.8}$$

non $s_0 = s(0)$, $s_1 = 1 - s_0$ eta $t_0 \geq 0$, $n(W)$ atarira iristen den arte. Ohartu $s_0 = 0$ dela $m(0) = m \leq n(W)$ bada eta $s_0 = 1$ bestela. $J^{n(W)}(m, W)$ funtzioa bi gaien batura bezala ikus daiteke, non lehenengo gaia m-n hasi eta t_0-raino doan fasea den, t_0 ataria ikutzen den unea delarik. Fase honetan kontrola s_0 da. Behin atarira iritsi denean sistema, kontrolen arteko trukatze bat gertaten da eta beraz bigarren gaia t_0 denboratik oreka ikutzen den arteko faseari dagokio. Bigarren fase honetan kontrola s_1 da. Izan ere, atari-politikek gehienez trukatze puntu bat dute, ikusi 4.5. Irudia.

Orain $\partial J^{n(W)}(m, W)/\partial m$ kalkulatuko da. $m \leq n(W)$ bada

$$\frac{\partial J^{n(W)}(m, W)}{\partial m} = \frac{EC^*(W) - C(m, 0) + W}{f^0(m)}, \tag{4.4.9}$$



Irudia 4.5: $n(W)$ politika: goian 1. Kasua, erdian 2. Kasua eta behean 3. Kasua. Gezi lodiek aukeratu den norabidea adierazten dute kasu bakoitzerako.

lortzen da eta $m > n(W)$ bada

$$\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W) - C(m,1)}{f^1(m)}. \tag{4.4.10}$$

Azken hau horrela ikus daiteke. Demagun $m \leq n(W)$, zeinak $s_0 = 0$ eta $s_1 = 1$ inplikatzen duen, orduan (4.4.8) ekuaziotik

$$\begin{aligned}\frac{\partial J^{n(W)}(m,W)}{\partial m} =& \frac{\mathrm{d}t_0}{\mathrm{d}m}\Big(C(n(W),0) - W - EC^*(W) - C(n(W),1) + EC^*(W)\Big) \\ &+ \int_0^{t_0} \frac{\mathrm{d}C(m^{n(W)}(t),0)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}m}\mathrm{d}t + \int_{t_0}^{\infty} \frac{\mathrm{d}C(m^{n(W)}(t),1)}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}m}\mathrm{d}t.\end{aligned} \tag{4.4.11}$$

$n(W)$ atari-politikak $\mathrm{d}m^{n(W)}(t)/\mathrm{d}t = f^0(m^{n(W)}(t))$ inplikatzen du edozein $t \in [0, t_0]$ tartean. Beraz,

$$\int_m^{n(W)} \frac{\mathrm{d}m^{n(W)}(t)}{f^0(m^{n(W)}(t))} = t_0,$$

zeinak

$$\begin{aligned}\frac{\mathrm{d}t_0}{\mathrm{d}m} &= \frac{\mathrm{d}}{\mathrm{d}m}\int_m^{n(W)} \frac{\mathrm{d}m^{n(W)}(t)}{f^0(m^{n(W)}(t))} = -\frac{1}{f^0(m)} + \int_m^{n(W)} \frac{\mathrm{d}}{\mathrm{d}m}\left(\frac{1}{f^0(m^{n(W)}(t))}\right)\mathrm{d}m^{n(W)}(t) \\ &= -\frac{1}{f^0(m)},\end{aligned} \tag{4.4.12}$$

inplikatzen duen. Izan ere, $\mathrm{d}(1/f^0(m^{n(W)}(t)))/\mathrm{d}m = 0$. $\mathrm{d}m^{n(W)}(t)/\mathrm{d}t = f^0(m^{n(W)}(t))$ denez $\mathrm{d}m/\mathrm{d}t = f^0(m)$ lortzen da. Azken hau eta (4.4.12) ekuazioa (4.4.11) ekuazioan ordezkatuz (4.4.9) lortzen da. (4.4.10) ekuazioa antzeko moduan lor daiteke.

Edozein $m \leq n(W)$ denean, $n(W)$ politikapean akzio optimoa *bandit*-a pasibo mantentzea da. Gainera, $\frac{\partial J^{n(W)}(m,W)}{\partial m}$ (4.4.6) ekuazioan ordezkatzean $\mathcal{J}_0(m,W) = 0$ lortzen da. $n(W)$ atari-politikak HJB (4.4.5) ekuazioa bete dezan, $\mathcal{J}_1(m,W) \geq 0$ betetzen dela frogatu behar da. $\frac{\partial J^{n(W)}(m,W)}{\partial m}$ (4.4.7) ekuazioan ordezkatuz hau

$$\mathcal{J}_1(m,W) \geq 0 \Leftrightarrow W \geq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}(C(m,1) - EC^*(W)), \tag{4.4.13}$$

ekuazioaren baliokidea dela ikus daiteke edozein $m \notin [m^1, m^0]$ eta $m \leq n(W)$, eta

$$\mathcal{J}_1(m,W) \geq 0 \Leftrightarrow W \leq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}(C(m,1) - EC^*(W)), \tag{4.4.14}$$

edozein $m \in [m^1, m^0]$ non $m \leq n(W)$. (4.4.13) ekuazioa edozein $m \notin [m^1, m^0]$ tarterako betetzen bada eta (4.4.14) ekuazioa edozein $m \in [m^1, m^0]$, tarterako orduan $n(W)$ politikapeko soluzioa *bandit*-a pasibo mantentzea, optimoa da.

Demagun $m > n(W)$. Beraz, $n(W)$ peko akzioa *bandit*-a aktibo uztea da. $\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W) - C(m,1)}{f^1(m)}$ (4.4.7) ekuazioan ordezkatuz $\mathcal{J}_1(m,W) = 0$ dela lortzen da. $n(W)$ politikak HJB (4.4.5) ekuazioa bete dezan, $\mathcal{J}_0(m,W) \geq 0$ dela frogatu behar da. $\frac{\partial J^{n(W)}(m,W)}{\partial m} = \frac{EC^*(W) - C(m,1)}{f^1(m)}$ ekuazioa (4.4.6) ekuazioan ordezkatuz,

$$\mathcal{J}_0(m,W) \geq 0 \Leftrightarrow W \leq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}(C(m,1) - EC^*(W)), \tag{4.4.15}$$

espresioaren baliokidea dela ikusten da edozein $m > n(W)$ bada. (4.4.15) baldintza edozein $m > n(W)$-rako betetzen bada orduan $n(W)$ politikapeko akzioa, *bandit*-a pasibo uztea, optimoa da.

Beraz, (4.4.13)−(4.4.15) baldintzak betetzen badira, orduan $n(W)$ politika optimoa da. Orduan (4.4.13)−(4.4.15) baldintzak betetzen direla ikustea falta da. Hau frogapenaren gainerantzekoan egingo da hiru kasu ezberdinetarako.

Lehenik eta behin $m^* = m^1$ onartuko da eta $W \leq w(m^1)$, hau da, 1. Kasua (2. Pausuan frogatu den legez). Beraz, $EC^*(W) = C(m^1, 1)$. Gogoratu $n(W)$ atari-politikak $W \geq w(m)$ inplikatzen duela edozein $m \leq n(W)$ bada eta $W \leq w(m)$ $m \geq n(W)$ bada. Beraz, $W \leq w(m^1)$ eta $w(m)$ ez-beherakorra izateak $n(W) \leq m^1$ inplikatzen du, ikusi 4.5. Irudia. (4.4.13)−(4.4.15) baldintzak hurrengo moduan idatz daitezke orduan: HJB beteko da baldin eta soilik baldin $W \geq (\leq) C(m,0) - C(m,1) + w^{(1)}(m)$ edozein $m \leq (\geq) n(W)$ bada. Azken hau $W \geq (\leq) w(m)$ edozein $m \leq (\geq) n(W)$-rako izatearen baliokidea da, izan ere, $w^{(1)}(m)$ ez-beherakorra da eta $W \leq w(m^1)$. Beraz, 1. Kasuan $n(W)$ politikak HJB ekuazioa betetzen du eta beraz optimoa da.

Antzeko moduan, $m^* = m^0$ bada eta $W \geq w(m^0)$, hau da, 3. Kasua, orduan $EC^*(W) = C(m^0, 0) - W$. $n(W)$ politikapean $W \geq w(m)$ denez edozein $m \leq n(W)$ denean eta $W \leq w(m)$ edozein $m \geq n(W)$ denean, orduan $w(m)$ ez-beherakorra izateak $n(W) \geq m^0$ inplikatzen du. Ikusi 4.5. Irudia. $EC^*(W) = C(m^0, 0) - W$ erabiliz, (4.4.13)−(4.4.15) baldintzak hurrengora sinplifikatzen direla lortzen da: $W \geq (\leq) C(m,0) - C(m,1) + w^{(0)}(m)$, edozein $m \leq (\geq) n(W)$ denean. Azken hau $W \geq (\leq) w(m)$ edozein $m \leq (\geq) n(W)$ denean izatearen baliokidea da, izan ere $w^{(0)}(m)$ ez-beherakorra da eta $W \geq w(m^0)$. Beraz, 3. Kasuan, $n(W)$ politikak HJB betetzen du eta beraz optimoa da.

Azkenik, 2. Kasua falta da, zeinetzat W-ren balioak $\frac{\mathrm{d}E(s^*,W)}{\mathrm{d}s^*} = 0$ betetzen duen, eta $s^* \in [0,1]$, hau da, $W = w(m^*)$. Beraz $n(W) = m^*$, ikusi 4.5. Irudia $n(W)$-ren definizioz. Testuinguru honetan

$$EC^*(W) = (1-s^*)(C(m^*,0) - W) + s^*C(m^*,1),$$

lortzen da. Azken hau (4.4.13) eta (4.4.15) baldintzetan ordezkatuz baldintza

$$W \geq (\leq) C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}\Big(C(m,1) - (1-s^*)(C(m^*,0) - W) - s^*C(m^*,1)\Big), \tag{4.4.16}$$

betetzera sinplifika daiteke, edozein $m \leq m^1(\ m \geq m^0)$ denean.

(4.4.14) eta (4.4.15) baldintzak

$$W \leq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}\Big(C(m,1) - (1-s^*)(C(m^*,0) - W) - s^*C(m^*,1)\Big), \tag{4.4.17}$$

idatz daitezke edozein $m \in [m^1, m^*]$ denean eta

$$W \leq C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}\Big(C(m,1) - (1-s^*)(C(m^*,0) - W) - s^*C(m^*,1)\Big), \tag{4.4.18}$$

edozein $m \in [m^*, m^0]$ tartean.

$f^1(m) \geq 0$ dela kontutan hartuz, edozein $m < m^1$ denean, eta $f^1(m) \leq 0$, bestela, eta $f^0(m)(1 - s^*) + s^*f^1(m) > 0$ hipotesitik edozein $m < m^*$ denean, eta $f^0(m)(1-s^*) + s^*f^1(m) < 0$ edozein $m^* > m$ denean, (4.4.16)–(4.4.18) baldintzak

$$\begin{aligned} W \geq (\leq) &\Bigg(C(m,0) - C(m,1) + \frac{f^1(m) - f^0(m)}{f^1(m)}\Big(C(m,1) - (1-s^*)C(m^*,0) - s^*C(m^*,1)\Big)\Bigg) \\ &\cdot \frac{f^1(m)}{f^0(m)(1-s^*) + s^*f^1(m)}, \end{aligned} \tag{4.4.19}$$

idatz daitezke edozein $m < m^*(m > m^*)$ denean. $W = w(m^*)$ denez, eta $w^{(2)}(\cdot)$ eta $w(\cdot)$ ez-beherakorrak direnez, (4.4.19) frogatzeko nahikoa da hurrengoa ikustea:

- RHS in (4.4.19)$\xrightarrow{m \to m^*} C(m^*,0) - C(m^*,1) + w^{(2)}(m^*)$,
- RHS in (4.4.19) $\leq C(m^*,0) - C(m^*,1) + w^{(2)}(m^*)$ edozein $m < m^*$ denean,
- RHS in (4.4.19) $\geq C(m^*,0) - C(m^*,1) + w^{(2)}(m^*)$ edozein $m > m^*$ denean.

Baldintza hauek 4.2. Leman frogatu dira (4.4.2. Eranskinean) eta beraz honek frogapena amaitzen du.

4.4.2 4.2. Lema

4.2 Lema. *Demagun $m_k^0 > m_k^1$. Izan bedi $(m_k^*, s_k^*) \in [m_k^1, m_k^0] \times [0,1]$ (4.1.3) problemaren oreka puntu optimoa. Definitu*

$$\begin{aligned} F_k(m_k) :=& \frac{f_k^1(m_k)}{f_k^0(m_k)(1-s_k^*) + s_k^* f_k^1(m_k)} \Bigg(C_k(m_k,0) - C_k(m_k,1) \\ &+ \frac{f_k^1(m_k) - f_k^0(m_k)}{f_k^1(m_k)} \Big(C_k(m_k,1) - (1-s_k^*) C_k(m_k^*,0) - s_k^* C_k(m_k^*,1) \Big) \Bigg), \end{aligned}$$

edozein m_k-rako. Orduan,

$$\lim_{m_k \to m_k^*} F_k(m_k) = C_k(m_k^*,0) - C_k(m_k^*,1) + w_k^{(2)}(m_k^*) = w_k(m_k^*), \tag{4.4.20}$$

$$F_k(m_k) \leq C_k(m_k^*,0) - C_k(m_k^*,1) + w_k^{(2)}(m_k^*), \text{ edozein } m_k < m_k^*, \tag{4.4.21}$$

$$F_k(m_k) \geq C_k(m_k^*,0) - C_k(m_k^*,1) + w_k^{(2)}(m_k^*) \text{ edozein } m_k > m_k^*. \tag{4.4.22}$$

Frogapena. 4.4.2. Sekzioan (4.4.20) frogatu da, eta 4.4.2. Sekzioan (4.4.21) eta (4.4.22) frogatu dira.

Frogapenean zehar k-rekiko dependentzia alde batera utziko da.

(4.4.20) **Ekuazioaren frogapena**

Kalkulu batzuen ostean $F(m)$

$$\frac{f^1(m)C(m,0) - f^0(m)C(m,1)}{f^1(m)s^* + f^0(m)(1-s^*)} - \frac{(f^1(m) - f^0(m))(1-s^*)C(m^*,0)}{f^1(m)s^* + f^0(m)(1-s^*)} - \frac{(f^1(m) - f^0(m))s^* C(m^*,1)}{f^1(m)s^* + f^0(m)(1-s^*)},$$

dela ikus daiteke zeinak

$$C(m^*,0) - C(m^*,1) + \frac{f^1(m)(C(m,0) - C(m^*,0))}{f^1(m)s^* + f^0(m)(1-s^*)} + \frac{f^0(m)(C(m^*,1) - C(m,1))}{f^1(m)s^* + f^0(m)(1-s^*)}, \tag{4.4.23}$$

inplikatzen duen. Bestalde,

$$\begin{aligned} &\lim_{m \to m^*} (f^1(m) - f^0(m)) \frac{(f^1(m)s^* + f^0(m)(1-s^*))}{m^* - m} = \left(\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*} s^* + \frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}(1-s^*) \right) (f^1(m^*) - f^0(m^*)) \\ &= -\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*} s^* f^1(m^*) + \frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*} s^* f^0(m^*) - \frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}(1-s^*) f^1(m^*) + \frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}(1-s^*) f^0(m^*), \end{aligned}$$

da, non lehenengo pausuan $s^* f^1(m^*) = (s^*-1) f^0(m^*)$ dela erabili den eta L'Hopital-en erregela erabili den. Lehenengo eta laugarrengo gaietan $s^* f^1(m^*) = (s^*-1) f^0(m^*)$ ordezkatuz

$$\lim_{m \to m^*} (f^1(m) - f^0(m)) \frac{(f^1(m)s^* + f^0(m)(1-s^*))}{m^* - m} = \frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*} f^0(m^*) - \frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*} f^1(m^*),$$

lortzen da. Azken espresio hau (4.4.23) ekuazioan ordezkatzen bada, non $(f^1(m) - f^0(m))/(m^* - m)$ espresioz frakzioak biderkatu eta zatitzen badira, $m \to m^*$ den heinean (4.4.23) Ekuazioa

$$
\begin{aligned}
&C(m^*,0) - C(m^*,1) + \frac{(f^1(m^*) - f^0(m^*))(f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*})}{f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*} - \frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}f^1(m^*)}, \\
&= C(m^*,0) - C(m^*,1) + w^{(2)}(m^*),
\end{aligned}
$$

lortzen da, izan ere, $\lim_{m\to m^*} \frac{C(m,a)-C(m^*,a)}{m^*-m} = -\frac{\mathrm{d}C(m^*,a)}{\mathrm{d}m^*}$, $a = 0, 1$ bada. Honek (4.4.20) ekuazioaren frogapena amaitzen du.

(4.4.21) eta (4.4.22) ekuazioen frogapena

$F(m) \leq (\geq) C(m^*,0) - C(m^*,1) + w^{(2)}(m^*)$ dela frogatu nahi da edozein $m < m^* (m > m^*)$ bada. $w^{(2)}(m^*)$ (4.4.23) ekuazioan ordezkatuz

$$
\begin{aligned}
&\frac{f^1(m)(C(m,0) - C(m^*,0))}{f^1(m)s^* + f^0(m)(1-s^*)} + \frac{f^0(m)(C(m^*,1) - C(m,1))}{f^1(m)s^* + f^0(m)(1-s^*)} \\
&\leq (\geq) \frac{(f^1(m^*) - f^0(m^*))(f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*})}{f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*} - \frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}f^1(m^*)},
\end{aligned}
\tag{4.4.24}
$$

lortzen da edozein $m < (>) m^*$ bada. Azken honetan $s^* = f^0(m^*)/(f^0(m^*) - f^1(m^*)) \geq 0$ ordezkatuko da, eta bi aldeak $f^0(m^*) - f^1(m^*)$ kenketaz zatituko dira, zeina gogoratu positiboa den $f^1(m^*) \leq 0$ eta $f^0(m^*) \geq 0$ baitira. Ez hori bakarrik, $f^a(\cdot)$ ez-gorakorra denaren hipotesia dela-eta $\bar{s}(\bar{m})$ hertsiki monotonoa $\nexists\, \bar{m} \in [m^1, m^0]$ zeinentzat $f^0(\bar{m}) = f^1(\bar{m}) = 0$ den. Orduan, (4.4.24) inekuazioa

$$
\begin{aligned}
&\frac{f^1(m)\,(C(m,0) - C(m^*,0))}{f^0(m^*)(f^1(m) - f^1(m^*)) - f^1(m^*)(f^0(m) - f^0(m^*))} + \frac{f^0(m)\,(C(m^*,1) - C(m,1))}{f^0(m^*)(f^1(m) - f^1(m^*)) - f^1(m^*)(f^0(m) - f^0(m^*))} \\
&\leq (\geq) \frac{f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*} - f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}}{f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*} - f^1(m^*)\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*}},
\end{aligned}
$$

idatz daiteke edozein $m < (>) m^*$ bada. Ezkerraldean $m - m^*$ kenketaz bidertatu eta zatituz

$$
\begin{aligned}
&\frac{-f^1(m)\left(\frac{C(m,0)-C(m^*,0)}{m-m^*}\right)}{f^1(m^*)\frac{f^0(m)-f^0(m^*)}{m-m^*} - f^0(m^*)\frac{f^1(m)-f^1(m^*)}{m-m^*}} + \frac{f^0(m)\left(\frac{C(m,1)-C(m^*,1)}{m-m^*}\right)}{f^1(m^*)\frac{f^0(m)-f^0(m^*)}{m-m^*} - f^0(m^*)\frac{f^1(m)-f^1(m^*)}{m-m^*}} \\
&\leq (\geq) \frac{f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*}}{f^1(m^*)\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*} - f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}},
\end{aligned}
\tag{4.4.25}
$$

lortzen da edozein $m < (>) m^*$ bada. $f^a(\cdot)$ funtzioa ganbila eta ez-gorakorra da $a = 0, 1$ bada. Orduan,

$$
\begin{aligned}
\frac{f^1(m) - f^1(m^*)}{m - m^*} &\leq (\geq) \frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*}, \text{ eta} \\
\frac{f^0(m) - f^0(m^*)}{m - m^*} &\leq (\geq) \frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*},
\end{aligned}
$$

edozein $m < (>) m^*$ bada. $f^1(m^*) \leq 0$ denez eta $f^0(m^*) \geq 0$

$$f^1(m^*)\frac{f^0(m) - f^0(m^*)}{m - m^*} - f^0(m^*)\frac{f^1(m) - f^1(m^*)}{m - m^*} \geq (\leq) f^1(m^*)\frac{\mathrm{d}f^0(m^*)}{\mathrm{d}m^*} - f^0(m^*)\frac{\mathrm{d}f^1(m^*)}{\mathrm{d}m^*},$$

lortzen da edozein $m < (>) m^*$, hau da, (4.4.25) ekuazioaren zenbakitzailea ordenaturik dago. Ohartu baita bi zenbakitzaileak hertsiki positiboak direla. Orain (4.4.25) ekuazioko izendatzaileek ere orden berdina jarraitzen dutela frogatuko da, hau da,

$$\begin{aligned}
&- f^1(m)\left(\frac{C(m,0) - C(m^*,0)}{m - m^*}\right) + f^0(m)\left(\frac{C(m,1) - C(m^*,1)}{m - m^*}\right) \\
&\leq (\geq) f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*},
\end{aligned} \tag{4.4.26}$$

edozein $m < m^*(m > m^*)$ denean. Honek frogapena amaitzen du.

Lehenik eta behin $m \in [m^1, m^*)(m \in (m^*, m^0])$ dela onartuko da. Orduan

$$\begin{aligned}
&- f^1(m)\left(\frac{C(m,0) - C(m^*,0)}{m - m^*}\right) + f^0(m)\left(\frac{C(m,1) - C(m^*,1)}{m - m^*}\right) \\
&\leq (\geq) - f^1(m)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*} + f^0(m)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*} \leq (\geq) - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*} + f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*},
\end{aligned}$$

edozein $m \in [m^1, m^*](m \in [m^*, m^0])$ bada. Lehenengo inekuazioa $C(\cdot, a)$ ganbila eta ez-beherakorra izateak inplikatzen du eta beraz $\frac{C(m,a)-C(m^*,a)}{m-m^*} \leq (\geq) \frac{\mathrm{d}C(m^*,a)}{\mathrm{d}m^*}$ $a = 0,1$ bada eta $m < (>) m^*$, eta $-f^1(m), f^0(m) \geq 0$ edozein m bada. Bigarren inekuazioa $\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*} \geq \frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}$ eta $f^1(m) - f^1(m^*) \geq (\leq) f^0(m) - f^0(m^*)$ espresioetatik ondorioztatzen da edozein $m < (>) m^*$ bada.

Orain suposatu $m \leq m^1$. Orduan

$$\begin{aligned}
&- f^1(m)\left(\frac{C(m,0) - C(m^*,0)}{m - m^*}\right) + f^0(m)\left(\frac{C(m^*,1) - C(m,1)}{m - m^*}\right) \\
&\leq (f^0(m) - f^1(m))\left(\frac{C(m,1) - C(m^*,1)}{m - m^*}\right) \leq (f^0(m^*) - f^1(m^*))\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m} \\
&\leq f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*},
\end{aligned}$$

edozein $m \leq m^1$ bada, non lehenengo inekuazioa $\left(\frac{C(m,0)-C(m^*,0)}{m-m^*}\right) \geq \left(\frac{C(m,1)-C(m^*,1)}{m-m^*}\right)$ edozein $m \leq m^1 \leq m^*$-rako eta $-f^1(m) \leq 0$ for all $m \leq m^1$ espresioek inplikatzen dute. Bigarren inekuazioa $f^0(m) - f^0(m^*) \leq f^1(m) - f^1(m^*)$ edozein $m \leq m^1$ eta $\frac{C(m,1)-C(m^*,1)}{m-m^*} \leq \frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m^*}$ edozein $m \leq m^*$-rako espresioek inplikatzen dute. Azkenik, hirugarren inekuazioa $\mathrm{d}C(m^*,0)/\mathrm{d}m^* \geq \mathrm{d}C(m^*,1)/\mathrm{d}m^*$ eta $-f^1(m^*) \geq 0$ espresioek inplikatzen dute.

Antzera

$$- f^1(m)\left(\frac{C(m,0) - C(m^*,0)}{m - m^*}\right) + f^0(m)\left(\frac{C(m,1) - C(m^*,1)}{m - m^*}\right) \geq f^0(m^*)\frac{\mathrm{d}C(m^*,1)}{\mathrm{d}m} - f^1(m^*)\frac{\mathrm{d}C(m^*,0)}{\mathrm{d}m^*},$$

da edozein $m \geq m^0$-rentzat, eta $f^0(m) < 0$, dela erabiliz $m \geq m^0$ denean eta $\left(\frac{C(m,0)-C(m^*,0)}{m-m^*}\right) \geq \left(\frac{C(m,1)-C(m^*,1)}{m-m^*}\right)$ edozein $m \geq m^0 \geq m^*$ denean. Beraz, (4.4.26) betetzen da edozein m-rako zeinak frogapena amaitzen duen. $\square$

4.4.3 4.2. Proposizioaren frogapena

k-ren gaineko dependentzia alde batera utziko da frogapen honetan. $\lambda \downarrow 0$ doan heinean indizea kalkulatuko da 0-1 motako atari-politikentzat lehenik eta 1-0 motako atari-politikentzat gero.

0-1 motako atari-politikak. Demagun (2.3.3) problema ebazten duen soluzio optimoa 0-1 motako atari-politika dela. Gogoratu $b^a(m) = \lambda\gamma$ dela hipotesiz eta $a = 0, 1$. Notazioa arintzearren hurrengoa definituko da frogapenean zehar.

$$\overline{D}^n(m) := \begin{cases} d^0(1) \cdot d^0(2) \cdot \ldots \cdot d^0(m), & \forall m \leq n, \\ d^0(1) \cdot \ldots \cdot d^0(n) \cdot d^1(n+1) \cdot \ldots \cdot d^1(m), & \forall m > n. \end{cases} \tag{4.4.27}$$

Ohartu $\overline{D}^n(m) = \overline{D}^{n-1}(m)$ edozein $m \leq n-1$ bada. Ohar hau frogapenean zehar erabiliko da.

Jaiotza-eta-heriotza motako prozesuak kontsideratu dira eta beraz $\pi^n(m) := \lambda^m\gamma^m\pi^n(0)/\overline{D}^n(m)$, non $\pi^n(0) = \left(\sum_{m=0}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^n(m)}\right)^{-1}$. Orain $\pi^{n-1}(0)/\pi^n(0)$ espresioa kalkulatuko da, zeina frogapenean erabiliko da geroaro. Ohartu

$$\frac{\pi^{n-1}(0)}{\pi^n(0)} = \frac{\sum_{m=0}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}{\sum_{m=0}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}} = 1 + \frac{\sum_{m=n}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^n(m)} - \sum_{m=n}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}}{\sum_{m=0}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}}, \tag{4.4.28}$$

dela, azken hau $\lambda \to 0$ bada $1 + \lim_{\lambda\downarrow 0} \frac{\mathcal{O}(\lambda^n)}{1+\mathcal{O}(\lambda)} = 1$ idatz daiteke.

Bestalde, $W(n)$ (2.3.6) ekuazioak definitzen du eta

$$W(n) = \frac{\xi_1(n) + \xi_2(n) + \xi_3(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))}, \tag{4.4.29}$$

idatz daiteke, non

$$\begin{aligned} \xi_1(n) &:= \sum_{m=0}^{n-1} C(m,0)(\pi^n(m) - \pi^{n-1}(m)), \\ \xi_2(n) &:= C(n,0)\pi^n(n) - C(n,1)\pi^{n-1}(n), \\ \xi_3(n) &:= \sum_{m=n+1}^{\infty} C(m,1)(\pi^n(m) - \pi^{n-1}(m)). \end{aligned} \tag{4.4.30}$$

Lehenik eta behin lehenengo gaia aztertuko da. Hau da,

$$\begin{aligned} \frac{\xi_1(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} &= \frac{\sum_{m=0}^{n-1} C(m,0)\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}(\pi^n(0) - \pi^{n-1}(0))}{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)}\pi^n(0) + (\pi^n(0) - \pi^{n-1}(0))\sum_{m=0}^{n-1} \frac{\lambda^m\gamma^m}{\overline{D}^n(m)}} \\ &= \frac{\sum_{m=0}^{n-1} C(m,0)\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)}\frac{1}{1-\frac{\pi^{n-1}(0)}{\pi^n(0)}} + \sum_{m=0}^{n-1} \frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}. \end{aligned} \tag{4.4.31}$$

(4.4.28) ekuazioan lortu den espresioa (4.4.31) ekuazioan ordezkatuko da. Orduan $\lambda \downarrow 0$ den heinean (4.4.31) ekuazioko zenbakitzailea

$$\lim_{\lambda\downarrow 0} - \frac{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)} \sum_{m=0}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}}{\sum_{m=n}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^n(m)} - \sum_{m=n}^{\infty} \frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}} + \sum_{m=0}^{n-1} \frac{\lambda^m\gamma^m}{\overline{D}^n(m)} = \lim_{\lambda\downarrow 0} - \frac{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)}(1+\mathcal{O}(\lambda))}{\left(\frac{\gamma^n}{\overline{D}^n(n)} - 1\right)(\lambda^n + \mathcal{O}(\lambda^{n+1}))} + 1 + \mathcal{O}(\lambda)$$
$$= -\left(1 - \frac{d^0(n)}{d^1(n)}\right)^{-1} + 1 = \frac{d^0(n)}{d^0(n) - d^1(n)},$$

bilakatzen da. Azken hau (4.4.31) ekuazioan ordezkatuz eta $\lambda \downarrow 0$ doan heinean

$$\frac{\xi_1(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = C(0,0)\left(1 - \frac{d^1(n)}{d^0(n)}\right) + \mathcal{O}(\lambda), \tag{4.4.32}$$

lortzen da.

Orain bigarren gaia aztertuko da. Hau da,

$$\frac{\xi_2(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = \frac{C(n,0) - C(n,1)\frac{d^0(n)}{d^1(n)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}{1 + (1 - \frac{\pi^{n-1}(0)}{\pi^n(0)})\frac{\overline{D}^n(n)}{\lambda^n\gamma^n}\sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}. \tag{4.4.33}$$

(4.4.28) ekuazioa azken ekuazio honetan ordezkatuko da. $\lambda \downarrow 0$ doan heinean (4.4.33) ekuazioko zenbakitzailea

$$\lim_{\lambda\downarrow 0} 1 - \frac{\frac{\overline{D}^n(n)}{\lambda^n\gamma^n}\left(\sum_{m=n}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)} - \sum_{m=n}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}\right)}{\sum_{m=0}^{\infty}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}} \sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)} = \lim_{\lambda\downarrow 0} 1 - \frac{1 - \frac{d^0(n)}{d^1(n)} + \mathcal{O}(\lambda)}{1 + \mathcal{O}(\lambda)}(1 + \mathcal{O}(\lambda))$$
$$= \frac{d^0(n)}{d^1(n)},$$

idatz daiteke, eta azken hau (4.4.33) ekuazioan ordezkatuz eta $\lambda \downarrow 0$ doan heinean

$$\begin{aligned}\frac{\xi_2(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} &= C(n,0)\frac{d^1(n)}{d^0(n)} - C(n,1) \\ &= C(n,0) - C(n,1) + C(n,0)\left(\frac{d^1(n)}{d^0(n)} - 1\right).\end{aligned} \tag{4.4.34}$$

Frogapena bukatzeko hirugarren gaia aztertuko da orain, hau da,

$$\frac{\xi_3(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = \frac{\sum_{m=n+1}^{\infty} C(m,1)\left(\frac{\lambda^m\gamma^m}{\overline{D}^n(m)} - \frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}\frac{\pi^{n-1}(0)}{\pi^n(0)}\right)}{\frac{\lambda^n\gamma^n}{\overline{D}^n(n)} + (1 - \frac{\pi^{n-1}(0)}{\pi^n(0)})\sum_{m=0}^{n-1}\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}}.$$

Azken hau, $\lambda \downarrow 0$ doan heinean $\frac{\mathcal{O}(\lambda^{n+1})}{\mathcal{O}(\lambda)+1}$ idatz daiteke, eta beraz,

$$\lim_{\lambda\downarrow 0}\frac{\xi_3(n)}{\pi^n(n) + \sum_{m=0}^{n-1}(\pi^n(m) - \pi^{n-1}(m))} = 0.$$

Azken hau (4.4.32) eta (4.4.34) ekuazioekin batuz $W(n) = W^{LT}(n) + o(1)$, lortzen da $\lambda \downarrow 0$ doan heinean, non

$$W^{LT}(n) = C(n,0) - C(n,1) + (C(n,0) - C(0,0)) \left(\frac{d^1(n)}{d^0(n)} - 1 \right).$$

Honek frogapena amaitze du.

1-0 motako atari-politikak. Demagun (2.3.3) problemaren soluzio optimoa 1-0 motako atari-politika dela. Gogoratu $b^a(m) = \lambda\gamma a$ hipotesia $a = 0,1$ bada. Notazioa arintzearren hurrengoa definituko da frogapenean zehar.

$$\overline{D}^n(m) := \begin{cases} d^1(1) \cdot d^1(2) \cdot \ldots \cdot d^1(m), & \forall m \leq n, \\ d^1(1) \cdot \ldots \cdot d^1(n) \cdot d^0(n+1), & m = n+1, \end{cases} \tag{4.4.35}$$

atari-politika n dela suposatu denean. Ohartu $\overline{D}^n(m) = \overline{D}^{n-1}(m)$ edozein $m \leq n-1$ bada. Ohar hau frogapenean zehar erabiliko da.

Tesi honetan jaiotza-eta-heriotza prozesuak kontsideratu dira eta beraz

$$\pi^n(m) := \lambda^m \gamma^m \pi^n(0) / \overline{D}^n(m),$$

edozein $m \leq n+1$, eta 0 bestela, non $\pi^n(0) = \left(\sum_{m=0}^{n+1} \frac{\lambda^m \gamma^m}{\overline{D}^n(m)} \right)^{-1}$.

Bestalde, $W(n)$ indizea (2.3.6) ekuazioak definitzen du

$$W(n) = \frac{\xi_1(n) + \xi_2(n) + \xi_3(n)}{\pi^n(n+1) - \pi^{n-1}(n)}, \tag{4.4.36}$$

non

$$\begin{aligned} \xi_1(n) &:= \sum_{m=0}^{n-1} C(m,1)(\pi^n(m) - \pi^{n-1}(m)), \\ \xi_2(n) &:= C(n,1)\pi^n(n) - C(n,0)\pi^{n-1}(n), \\ \xi_3(n) &:= C(n+1,0)\pi^n(n+1), \end{aligned} \tag{4.4.37}$$

eta $\pi^n(m)$ m egoeraren oreka egoerako probabilitatea.

Orain $\pi^{n-1}(0)/\pi^n(0)$ espresioa kalkulatuko da, zeina frogapenean erabiliko den. Beraz

$$\begin{aligned} \frac{\pi^{n-1}(0)}{\pi^n(0)} &= \frac{\sum_{m=0}^{n+1} \frac{\lambda^m \gamma^m}{\overline{D}^n(m)}}{\sum_{m=0}^{n} \frac{\lambda^m \gamma^m}{\overline{D}^{n-1}(m)}} = 1 + \frac{\frac{\lambda^n \gamma^n}{\overline{D}^n(n)} + \frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^{n+1}(n+1)} - \frac{\lambda^n \gamma^n}{\overline{D}^{n-1}(n)}}{\sum_{m=0}^{n} \frac{\lambda^m \gamma^m}{\overline{D}^{n-1}(m)}} \\ &= 1 + \frac{\frac{\lambda^n \gamma^n}{\overline{D}^n(n-1)} \left(\frac{1}{d^1(n)} - \frac{1}{d^0(n)} \right) + \frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^{n+1}(n+1)}}{\sum_{m=0}^{n} \frac{\lambda^m \gamma^m}{\overline{D}^{n-1}(m)}}, \end{aligned} \tag{4.4.38}$$

da. Azken hau $\lambda \to 0$ doan heinean $1 + \lim_{\lambda \downarrow 0} \frac{\mathcal{O}(\lambda^n)}{1+\mathcal{O}(\lambda)} = 1$, bilakatzen da.

Lehenik eta behin lehenengo gaia aztertuko da. Hau da,

$$\frac{\xi_1(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = \frac{\sum_{m=0}^{n-1} C(m,1)\frac{\lambda^m\gamma^m}{\overline{D}^n(m)}\left(1-\frac{\pi^{n-1}(0)}{\pi^n(0)}\right)}{\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)} - \frac{\lambda^n\gamma^n}{\overline{D}^{n-1}(n)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}. \tag{4.4.39}$$

Orain (4.4.38) espresioa ordezkatuko da (4.4.39) ekuazioan. Orduan,

$$\begin{aligned}
&\frac{1-\frac{\pi^{n-1}(0)}{\pi^n(0)}}{\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)} - \frac{\lambda^n\gamma^n}{\overline{D}^{n-1}(n)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}\\
&= -\frac{\frac{\lambda\gamma}{d^0(n+1)d^1(n)} + \frac{1}{d^1(n)} - \frac{1}{d^0(n)}}{\sum_{m=0}^{n}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}} \cdot \frac{1}{\frac{\lambda\gamma}{d^0(n+1)d^1(n)} - \frac{1}{d^0(n)} - \frac{\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)} + \frac{\lambda^n\gamma^n}{\overline{D}^{n-1}(n-1)}\left(\frac{1}{d^1(n)}-\frac{1}{d^0(n)}\right)}{\sum_{m=0}^{n}\frac{\lambda^m\gamma^m}{\overline{D}^{n-1}(m)}}}\\
&= \frac{d^0(n)}{d^1(n)} - 1 + o(\lambda),
\end{aligned}$$

non azken inekuazioa $\lambda \to 0$ doan heinean lortzen den. Azken hau (4.4.39) ekuazioan ordezkatuz eta $\lambda \downarrow 0$ doan heinean

$$\frac{\xi_1(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = C(0,1)\left(\frac{d^0(n)}{d^1(n)} - 1\right) + \mathcal{O}(\lambda). \tag{4.4.40}$$

Orain bigarren gaia aztertuko da, hau da,

$$\frac{\xi_2(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = \frac{C(n,1)\frac{d^0(n)}{d^1(n)} - C(n,0)\frac{\pi^{n-1}(0)}{\pi^n(0)}}{\frac{\lambda\gamma d^0(n)}{d^0(n+1)d^1(n)} - \frac{\pi^{n-1}(0)}{\pi^n(0)}}. \tag{4.4.41}$$

(4.4.38) ekuazioa azken ekuazioa ordezkatuko da. $\lambda \downarrow 0$ doan heinean (4.4.41)

$$\frac{\xi_2(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = -C(n,1)\frac{d^0(n)}{d^1(n)} + C(n,0) + \mathcal{O}(\lambda),$$

bilakatzen da. Beraz,

$$\lim_{\lambda\downarrow 0}\frac{\xi_2(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = C(n,0) - C(n,1) + C(n,1)\left(1-\frac{d^0(n)}{d^1(n)}\right). \tag{4.4.42}$$

Frogapena bukatzearren hirugarren gaia aztertuko da, hau da,

$$\frac{\xi_3(n)}{\pi^n(n+1)-\pi^{n-1}(n)} = \frac{C(n+1,0)\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)}}{\frac{\lambda^{n+1}\gamma^{n+1}}{\overline{D}^n(n+1)} - \frac{\lambda^n}{\overline{D}^{n-1}(n)}\frac{\pi^{n-1}(0)}{\pi^n(0)}} = \frac{C(n+1,0)}{1-\frac{\overline{D}^n(n+1)}{\lambda\gamma\overline{D}^{n-1}(n)}\frac{\pi^{n-1}(0)}{\pi^n(0)}}.$$

Azken hau, $\lambda \downarrow 0$ doan heinean, $\frac{\mathcal{O}(\lambda)}{\mathcal{O}(\lambda)+1}$ bilakatzen da eta beraz

$$\lim_{\lambda \downarrow 0} \frac{\xi_3(n)}{\pi^n(n+1) - \pi^{n-1}(n)} = 0.$$

Azken hau (4.4.40) eta (4.4.42) ekuazioekin batuz $W(n) = W^{LT}(n) + o(1)$, lortzen da $\lambda \downarrow 0$ doan heinean, non

$$W^{LT}(n) = C(n,0) - C(n,1) + (C(n,1) - C(0,1)) \left(1 - \frac{d^0(n)}{d^1(n)}\right).$$

Zeinak frogapena amaitzen duen.

4.4.4 4.3. Proposizioaren frogapena

k-ren gaineko dependentzia alde batera utziko da frogapen honetan.

$d^a(m) > 0$ da $a = 0,1$ bada eta edozein $m > 0$ bada, eta $b^a(m) = \lambda\gamma$ da edozein m eta $a = 0,1$ badira eta 0-1 egiturako atari-politika batek (2.3.3) optimoki ebazten duenean. Antzeko moduan $d^a(m) > 0$ da edozein $m > 0$ eta $a = 0,1$ badira eta $b^a(m) = \lambda\gamma a$ da edozein m eta $a = 0,1$ badira eta 1-0 egiturako atari-politikak (2.3.3) optimoki ebazten dutenean.

$\lambda \downarrow 0$ den heinean $f^a(m) \to -d^a(m)$, eta hipotesiz $d^a(m) > 0$ edozein m-rako, orduan m-ren balioak $-d^a(m) = 0$ betetzen badu $m < 0$ da. Beraz, 4.1.2. Sekzioko hipotesia dela-eta $m^a = 0$ dela ondorioztatzen da $a = 0,1$ bada. Azken honek 4.1. Proposizioarekin batera indize fluidoa 0-1 egiturako politikentzat $w(m) = C(m,0) - C(m,1) + w^{(0)}(m)$ ondorioztatzen da edozein m-rako eta 1-0 egiturako politikentzat $w(m) = C(m,0) - C(m,1) + w^{(1)}(m)$ edozein m-rako $m^0 = m^1 = 0$ baita. Orduan, 4.2. Proposiziotik

$$\lim_{\lambda \downarrow 0} W(m) = C(m,0) - C(m,1) + \frac{C(m,a) - C(0,a)}{d^a(m)} \left(d^1(m) - d^0(m)\right) = \lim_{\lambda \downarrow 0} w(m).$$

Honek frogapena amaitzen du.

4.4.5 4.4. Proposizioaren frogapena

Frogapenean zehar k-rekiko dependentzia alde batera utziko da.

$n \to \infty$ doan heinean, indize fluidoa $w(n) = C(n,0) - C(n,1) + \delta(\mu + \theta') - \delta'\theta' + w^{(3)}(n)$ bilakatzen da. $C(n,a)$ goitik P eta Q mailako polinomioz bornatuta dagoela suposatu da $a = 0,1$ bada. Beraz, $C(n,a) = E(n,a) + o(1)$ eta $E(n,1) = \sum_{i=0}^{P} C^{(P,i)} n^i$ idatz daitezke n-ren balio handietarako, non

$$C^{(P,i)} := \lim_{n \to \infty} \frac{C(n,1) - \sum_{j=i+1}^{P} C^{(P,j)} n^j}{n^i},$$

eta $E(n,0) = \sum_{i=0}^{Q} E^{(Q,i)} n^i$, non

$$E^{(Q,i)} := \lim_{n \to \infty} \frac{C(n,0) - \sum_{j=i+1}^{Q} E^{(Q,j)} n^j}{n^i}.$$

Orduan, $n \to \infty$ den heinean $w(n) = w^{\infty}(n) + o(1)$, non $w^{\infty}(n) = \delta(\mu + \theta') - \delta'\theta' + w^c(n) + o(1)$, eta

$$w^c(n) = E(n,0) - E(n,1) + \frac{(\mu + \theta' - \theta)}{\theta} \frac{(E(n,0) - E(\lambda/\theta, 0))}{n - \lambda/\theta}.$$

Ohartu $(E(n,0) - E(\lambda/\theta, 0))/(n - \lambda/\theta)$ n-ren balio handietarako

$$\begin{aligned}
\frac{\sum_{i=0}^{Q} E^{(Q,i)} n^i - \sum_{i=0}^{Q} E^{(Q,i)} (\lambda/\theta)^i}{n - \lambda/\theta} &= \sum_{i=0}^{Q} E^{(Q,i)} \frac{(n^i - (\lambda/\theta)^i)}{n - \lambda/\theta} = \sum_{i=2}^{Q} E^{(Q,i)} \left(\sum_{j=0}^{i} \left(\frac{\lambda}{\theta}\right)^j n^{i-1-j} \right) \\
= \frac{E(n,0)}{n} + \frac{E^{(Q,1)} \left(\frac{\lambda}{\theta}\right) + E^{(Q,2)} \left(\frac{\lambda}{\theta}\right)^2 + \ldots + E^{(Q,Q)} \left(\frac{\lambda}{\theta}\right)^Q}{n} &+ \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \\
= \frac{E(n,0)}{n} + \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} + o(1), & \qquad (4.4.43)
\end{aligned}$$

ere idatz daitekeela. Orain $\lim_{n\to\infty} W(n)/w(n)$ kalkulatuko da, zeina (3.3.4) ekuazioko emaitza dela eta

$$\begin{aligned}
\lim_{n\to\infty} \frac{W(n)}{w(n)} &= \lim_{n\to\infty} \frac{W^{\infty}(n) + o(1)}{w^{\infty}(n) + o(1)} = \lim_{n\to\infty} \frac{\delta(\mu + \theta') - \delta'\theta' + W^c(n) + o(1)}{\delta(\mu + \theta') - \delta'\theta' + w^c(n) + o(1)} \\
&= \lim_{n\to\infty} \frac{E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta} \left(\frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \right) + \mathcal{O}(1)}{E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta} \left(\frac{E(n,0)}{n} + \sum_{i=2}^{Q} E^{(Q,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \right) + \mathcal{O}(1)} \\
&= 1 + o(1),
\end{aligned}$$

ekuazioaren baliokidea den. Azken hau zenbakitzailea eta izendatzailearen gairik handiena $E(n,0) - E(n,1) + \frac{(\mu+\theta'-\theta)}{\theta} \frac{E(n,0)}{n}$ ekuazioak emateak inplikatzen du. Honek (4.2.1) espresioaren frogapena amaitzen du.

Orain (4.2.2) espresioa lortuko da $P = Q$ onartuz eta $C^{(P,i)} = E^{(P,i)}$ edozein $i \in \{2, \ldots, P\}$ bada. Ohartu hipotesi honen pean (4.4.43) ekuazioa $(E(n,0) - E(\lambda/\theta, 0))/(n - \lambda/\theta)$ da n-ren balio handietarako eta

$$\frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} + o(1),$$

idatz daiteke. Beraz,

$$\begin{aligned}
w^{\infty}(n) =& \delta(\mu + \theta') - \delta'\theta' + E(n,0) - E(n,1) \\
&+ \frac{(\mu + \theta' - \theta)}{\theta} \left(\frac{E(n,0)}{n} + \sum_{i=2}^{P} C^{(P,i)} \sum_{j=0}^{i-2} n^{i-2-j} \left(\frac{\lambda}{\theta}\right)^{j+1} \right) + o(1).
\end{aligned}$$

(3.3.4) ekuazioko emaitza dela eta Then, by the result in (3.3.4) we have $W^{\infty}(n) = w^{\infty}(n) + o(1)$, lortzen da eta beraz, $W(n) = w(n) + o(1)$ n-ren balio handietarako, zeinak frogapena amaitzen duen.

5. Kapitulua

Kontrol fluido optimoa uzteak gerta daitezken ilara batean

Contents

5.1 Sarrera 106
5.2 Ereduaren deskribapena 107
5.3 Kontrol fluido eredua 108
5.4 Bezero klase kopuru arbitrarioa 114
5.5 Emaitza numerikoak 116
5.6 Eranskina 120

Kapitulu honetan kostu linaleko uzteak gerta daitezken ilara bat aztertuko da. Eredu hau 3. Kapituluan aztertu den ereduaren kasu partikular bat da non mantentze-kostu ganbilak kontsideratu diren. Bi egoera ezberdin izango dira aztergai: (1) bezeroak ilaran zain daudenean ala zerbitzatuak izaten ari direnean utzi dezakete sistema, (2) ilaran zain dauden bezeroek soilik utzi dezakete sistema. 3. Kapituluan erlaxazio Lagrangearra erabili da heuristika bat lortzeko. Kapitulu honetan kontrol eredu fluido bat aurkeztuko da determinista dena jatorrizko eredua hurbiltzeko. Lan karga handia denean $\tilde{c}\mu/\theta$ erregela dela eredu fluidoaren soluzio optimoa ikusiko da. Lan karga txikia denean Pontyagin-en minimoaren printzipioa erabiliko da eredu fluidoaren soluzio optimoa lorzeko bi bezero klase dauden kasuan: trukatze-funtzio bat existitzen da zeinak egoeren espazioa bitan banatzen duen. Trukatze-futnzioaren gainetik $\tilde{c}\mu/\theta$ handiena duen klaseak izango du lehentasuna eta trukatze-funtziotik behera $\tilde{c}\mu$ handiena duen bezero klaseak izango du lehentasuna. Egitura berbera ikusi da eredu estokastikoan bezero klase arbitrarioarekiko. Soluzio honetan oinarrituz heuristika bat aurkeztuko da eta analisi numerikoak erakutsiko du heuristika honen errendimendua ona dela lan karga ezberdinetarako.

5.1 Sarrera

Kapitulu honetan baliabide bakar bat bezero klase ezberdinen artean nola banatu aztertuko da uzteak gerta daitezkeen ilaretan. Sarreran eztabaidatu den legez, eta 3. Kapituluan azaldu denez, soluzio optimoa karakterizatzea oso konplikatua gerta daiteke. 3. Kapituluan erlaxazio Lagrangearra aplikatu da jatorrizko ereduarentzat heuristika bat lortzeko. Kapitulu honetan doiketa fluidoa aplikatuko da. Literaturan hainbat lan aurki daitezke uzteak gerta daitezkeen ilaren kontrola aztertzen dutenak mugako eremuetan. Hurbilketa bat Halfin-Whitt-en trafiko geldoko eremua da. Hau da, iritsiera tasa totala eta zerbitzari kopurua biak nahiko handiak direnean lortzen den eremua, trafikoaren intentsitateak 1-runtz konbergitzen duelarik. Uzteak gerta daitezkeen ilaren kasuan hurbilketa hau [44] artikuluan aztertu da. Doiketa honek difusio kontrol problema bat sortzen du, zeinarentzat kontrol optimoa [10], eta [54] artikuluetan aztertu diren eta egoerarekiko dependentzia dutela erakutsi da. Lan karga handiko eremuetan ilaren uzteak doiketa fluido bidez [8] eta [9] artikuloetan aztertu dira, non autoreek zerbitzari kopurua eta iritsiera tasa doitu dituzten $\tilde{c}\mu/\theta$ erregela lortzeko (*i.e.* lehentasuna $\tilde{c}\mu/\theta$ indiziaren araberakoa denean). Erregela hau fluido asintotikoki optimoa dela frogatu da. Beraien analisian lan karga handia izatearen hipotesia beharrezkoa da, izan ere, eredu fluidoaren traiektoriak hertsiki positiboa den egoera batera konbergitzen dute zeinak guztiz karakteritzatzen duen batez besteko errendimendua. $\tilde{c}\mu/\theta$ erregela naturalki dator *absorbing* egoeraren errendimendua optimizatzen duen politika gisa. Uzterik gabeko ilara batean, $\tilde{c}\mu$ erregela, *i.e.*, lehentasuna $\tilde{c}\mu$ indizearen araberakoa denean, optimoa da ilara zerbitzari bakar klase anitz batean batez besteko kostuak deskontatuak direnean, lehentasunezko eta lehentasunik gabeko kasuetan, ikusi [33] adibidez.

Hemen aztergai den eredu estokastikoa 3. Kapituloko eredua da mantentze-kostu linearretarako. Analisia bi eredu ezberdinetan egingo da: (1) bezeroak ilaran zein zerbitzua jasotzean utzi dezakete sistema, (2) ilaran zain dauden bezeroek utzi dezakete sistema baina zerbitzuan daudenek ez.

3. Kapituloan aurkeztu den notazioa erabiliko da hemen, non (1) kasua $\theta' = \theta$-ri dagokion eta (2) kasua $\theta' = 0$-ri.

Lan karga handia denean, oreka puntu optimoa karakterizatuko da eta $\tilde{c}\mu/\theta$ erregela aplikatuz dinamikak oreka puntu honetara konbergitzen duela frogatuko da, oreka puntu hau hertsiki positiboa da. Lan karga txikiko eremuan, eredu fluidoak sistema hustuko du denbora finituan. Azken honek analisia zailtzen du lan karga handien eremuarekiko alderatuz. PMP erabiliz soluzio optimoa karakterizatuko da bi bezero klaseko ilara baterako. Soluzio optimoak egitura harrigarria du, trukaze-funtzio bat existitzen da bi dimentsioko egoera-espazioa bitan banatzen duena: bezero kopurua nahiko txikia denean $\tilde{c}\mu$ erregela da optimoa eta bezero kopurua nahiko handia denean orduan $\tilde{c}\mu/\theta$ erregela da optimoa. Gogoratu 3. Kapituloan $\tilde{c}\mu/\theta$ erregela lortu dela. Numerikoki problema estokastikoa ebatziz ikusi ahal izan da trukatze-funtzioa bat existitzen dela jatorrizko ereduarentzat ere (lehentasuna $\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ inizeen araberakoa da), eta funtzio hau ongi hurbiltzen du sistema fluidoarentzat lortu den trukatze-funtzioarekin. Gainera, sistema estokastikoan bi baino bezero klase gehiago duden sistemetan ere $\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ erregelen arteko erlazio bat dagoela ikusi da. Intuizio hau erabiliz heuristika bat proposatu da eredu estokastikoarentzat klase kopurua arbitrarioa denean. Azkenik, fluidoan oinarritutako heuristika honen errendimendua aztertu da numerikoki eta ikusi da heuristika honek eragiten duen errore erlatiboa nahiko txikia dela. Kapitulu honetan proposatutako heuristikak errendimendu ona erakusten du lan karga askotarako, ordez $\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ indize politikak errendimendu ona dute lan karga txiki edo handietan soilik.

Kapitulu honen gainerakoak hurrengo egitura jarraitzen du. 5.2. Sekzioan uzteak gerta daitezken optimizazio problema eredu estokastikoa aurkeztuko da. 5.3. Sekzioan honen eredu fluidoa aurkeztuko da, lan karga txikietarako kasuan soluzio optimoa aurkituko da bi bezero klase daudenean sisteman eta lan karga handietan soluzio optimoa aurkituko da. 5.4. Sekzioan heuristika bat proposatuko da bezero klase arbitrariorako eta 5.5. Sekzioan numerikoki aztertuko da heuristika honen errendimendua eta hainbat indize politiken errendimenduarekin konparatuko da. Frogapen gehienak 5.6. Eranskinean aurki daitezke.

5.2 Ereduaren deskribapena

K bezero klaseko ilara zerbitzari-bakar klase-anitza da aztergai. k klaseko bezeroak λ_k tasako Poisson prozesu bat jarraituz iristen dira eta zerbitzuak μ_k tasako banaketa esponentziala jarraitzen du. k klaseko bezeroak sistema utzi dezake ilaran zain dagoenean θ_k tasako banaketa esponentziala duen denbora bat pasa ostean, eta θ'_k tasako banaketa esponentziala duen denbora baten ostean dagoeneko zerbitzua jasotzen ari bada. Izan bedi $\rho_k = \lambda_k/\mu_k$ k klaseko bezeroen lan karga eta $\rho = \sum_k \rho_k$ lan karga osoa. Zerbitzariaren edukiera 1 dela onartuko da eta gehienez bezero bat har dezake zerbitzura denbora unitateko, zerbitzua lehetasunezkoa izan daiteke. Uneoro ϕ politikak erabakiko du zein klase zerbitzatuko duen zerbitzariak. Markov propietatea dela eta egoerarekiko menpekotasuna duten politiketan zentratuko da analisia, non egoera bezero kopurua den. ϕ politika finko batentzat, $(S_1^\phi(t), \dots, S_K^\phi(t))$ kontrol aldagaiak, t denboran bezeroen klase bat zerbitzatua den ala ez adierazten du, *i.e.*, t denboran k klaseko bezero bat badago zerbitzuan, orduan $S_k^\phi(t) = 1$ eta $S_l^\phi(t) = 0$ edozein $l \neq k$ bada. Beraz, $S_k^\phi(t) \in \{0, 1\}$ eta $\sum_{k=1}^K S_k^\phi(t) \leq 1$ betetzen dira.

Bi eredu ezberdin izango dira aztergai, bezeroek sistema utzi dezaketen unearen arabera:

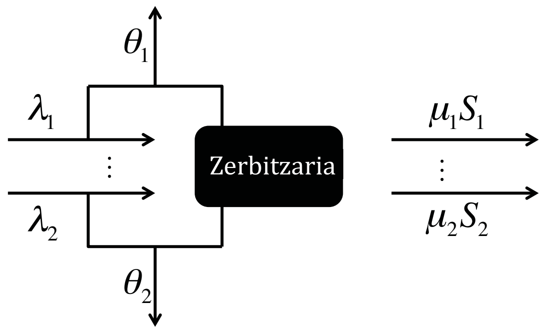
- $\theta = \theta'$ kasua: k klaseko bezeroak ilaran zain dagoenean zein zerbitzua jasotzen ari denean utzi dezake sistema, ikusi 5.1a. Irudia. Beraz, $\theta_k = \theta'_k$. Bestalde, $\delta_k = \delta'_k$ onartuko da edozein $k \in \{1, \dots, K\}$ bada.
- $\theta' = 0$ kasua: bezeroak ilaran zain dagoenean soilik utzi dezake sistema, ikusi 5.1b. Irudia. Beraz, $\theta'_k = 0$ edozein $k \in \{1, \dots, K\}$ bada.

Bi ereduak literaturan aztertuak izan dira *e.g.*, [40] artikuluan $\theta = \theta'$ eredua aztertu da, ordez [8, 9, 15] artikuluetako autoreek $\theta' = 0$ eredua aztertu dute. 3. Kapituluko analisian bi kasuak kontsideratu dira.

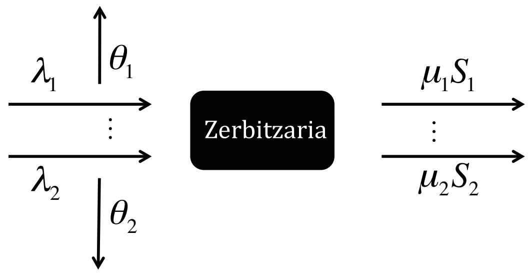
ϕ politika finko batentzat, $N_k^\phi(t)$ k klaseko sisteman dauden bezero kopuruaren adierazle den aldagaia da ($\theta = \theta'$ kasuan), edo ilaran dauden k klaseko bezero kopuruaren adierazle ($\theta' = 0$ kasuan). Izan bedi c_k k klaseko bezero bat sisteman mantentzeagatik ordaintzen den denbora unitateko kostua. Izan bedi δ_k k klaseko bezero bakoitzak uztean eragiten duen kostua. Helburua betez besteko kostua minimizatzea da, hau da,

$$\min_\phi \limsup_{T\to\infty} \sum_{k=1}^K \frac{1}{T}\mathbb{E}\left(\int_0^T c_k N_k^\phi(t)\ \mathrm{d}t + \delta_k R_k^\phi(T)\right) = \min_\phi \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T \sum_{k=1}^K \tilde{c}_k N_k^\phi(t)\ \mathrm{d}t\right),$$

non $R_k^\phi(T)$ $[0, T]$ tartean sistema utzi duten k klaseko bezeroen kopuruaren adierazle den, $\tilde{c}_k := c_k + \delta_k\theta_k$, $k = 1, \dots, K$, eta $\mathbb{E}(R_k^\phi(T)) = \theta_k\mathbb{E}(\int_0^T N_k^\phi(t)\mathrm{d}t)$. Ohartu $\theta' = 0$ kasuan, ilaran dauden bezeroek soilik utzi dezaketela sistema eta beraz uzteagatik sortutako kostuak eurek bakarrik eragin ditzakete. Bestalde, inplizituki $\theta' = 0$ kasuan, ilaran zain dauden bezeroek bakarrik eragiten dituztela mantentze-kostuak



(a) $\theta = \theta'$ kasua



(b) $\theta' = 0$ kasua

Irudia 5.1: Uzteak gerta daitezkeen ilara zerbitzari-bakar klase-anitza

onartu da, azken hau [8, 9] artikuluetan ere onartu da, ordez [15] artikuluan bezero guztiek eragiten dute mantentze-kostua. $\theta = \theta'$ kasuan, bezero guztiek eragiten dute uzte kostua eta baita mantentze-kostua ere.

Eredu hauek oso zailak gerta daitezke ebazteko, hau 3.1. Sekzio bukaeran eztabaidatu da.

5.3 Kontrol fluido eredua

Sekzio honetan 5.2. Sekzioan aurkeztu diren eredu estokastikoak ($\theta = \theta'$, $\theta' = 0$) eredu fluido determinista batez bidez hurbildu dira, non batez besteko dinamika bakarrik hartu den kontutan, ikusi 1.3.2. Sekzioa. Hau da, izan bedi $m_k(t)$ k klaseko fluido kantitatea eta $s_k(t)$ kontrol parametroa. Orduan dinamika fluidoa hurrengo ekuazio diferentzialek deskribatzen dute:

$$\frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \lambda_k - \mu_k s_k(t) - \theta_k m_k(t), \text{ edozein } k \in \{1, \ldots, K\} \text{ denean,}$$

$$(s_1(t), \ldots, s_K(t)) \in \mathcal{S}, \; m_k(t) \geq 0, \text{ edozein } k \in \{1, \ldots, K\}, \text{ eta edozein } t \text{ badira,}$$

eta

$$\mathcal{S} := \{s = (s_1, \ldots, s_K) \text{ s.t } \sum_{k=1}^{K} s_k \leq 1, \, s_k \geq 0, \text{ edozein } k \in \{1, \ldots, K\}\}.$$

Analisi fluidoa bi zatitan banatuko da: (1) $\rho < 1$, zeina karga txiki izenez adieraziko den, eta (2) $\rho > 1$ zeina karga handi izenez adieraziko den. Ohartu $\rho < 1$ den kasuan, edozein *non-idling* politikak (*i.e.*, $\sum_{k=1}^K s_k(t) = 1$, $\sum_{k=1}^K m_k(t) > 0$ bada) $(0, \ldots, 0)$[1] puntura konbergituko du. Beraz, $\rho < 1$ denean $(0, \ldots, 0)$ puntura iritsi arteko kostu totala minimizatzea izango da helburu. Azken hau hurrengo moduan idatz daiteke:

$$\min_{s(t) \in \mathcal{S}} \int_0^\infty \sum_{k=1}^{K} \tilde{c}_k m_k(t) \; \mathrm{d}t.$$

[1] Izan bedi $w(t) := \sum_{k=1}^K m_k(t)/\mu_k$. Orduan $\frac{\mathrm{d}w(t)}{\mathrm{d}t} = \rho - \sum_{k=1}^K s_k(t) - \sum_{k=1}^K \frac{\theta_k}{\mu_k} m_k(t) < \rho - 1 < 0$, beraz $w(t)$ funtzioak zerorantz konbergitzen du.

$\rho > 1$ demean, oreka puntua $(0, \ldots, 0)$ ez beste puntu bat izango da. Beraz, $\rho > 1$ kasuan helburua batezbesteko kostua minimizatzea izango da, hau da,

$$\min_{s(t)\in\mathcal{S}} \lim_{T\to\infty} \frac{1}{T} \int_0^T \sum_{k=1}^{K} \tilde{c}_k m_k(t)\ \mathrm{d}t.$$

Kapitulu honetan zehar problema hau P Problema izenez adieraziko da.

5.3.1 Lan karga txikian politika optimoa bi bezero klaserako

Sekzio honetan $\rho < 1$ kasua aztertuko da eta kontrol fluido eredua ebatziko da. Bi bezero klase direneko kasuan zentratuko da analisia, zeinarentzat soluzioa nahiko konplikatua den. Halere, soluzio honek intuizioa eskeintzen du bezero klase kopurua arbitrario den kasurako politika bat definitzeko.

Soluzio optimoak bi egitura posible dituela ikusiko da: alde batetik, trukatze-funtzio bat ager daiteke, *i.e.*, klase bati eskeiniko zaio lehentasuna trukatze-funtzioaren gainetik eta beste klaseari azpitik, beste alde batetik, lehentasuna egoera-espazio guztian klase berdinari emango zaio. Honek lau politika ezberdin ahalbideratzen ditu. Hurrengo proposizioan ikusiko den legez, soluzio optimoa $\tilde{c}_1\mu_1$ eta $\tilde{c}_2\mu_2$, eta $\tilde{c}_1\mu_1/\theta_1$ eta $\tilde{c}_2\mu_2/\theta_2$ indizeen ordenak karakterizatzen du. Frogapena 5.6.2. Eranskinean aurki daiteke.

5.1 Proposizioa. *Demagun $K = 2$ eta izan bedi $\lambda_k, \mu_k, \theta_k, c_k$ eta δ_k finkoak edozein $k \in \{1, 2\}$ bada. Demagun $\rho < 1$. $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$ bada, orduan P problemaren soluzio optimoa kostu totala minimizatzerakoan hurrengoa da:*

- *$\tilde{c}_2\mu_2 \leq \tilde{c}_1\mu_1$ bada, orduan*
  - *$s^* = (0, 1)$, $m_2 > h(m_1)$ bada,*
  - *$s^* = (1, 0)$, $m_2 \leq h(m_1)$ bada eta $m_1 > 0$,*
  - *$s^* = (\rho_1, 1 - \rho_1)$, $m_2 \leq h(0)$ bada $m_1 = 0$,*

  non $h(\cdot)$ trukatze-funtzioa

$$h(m_1) := \frac{a_1 m_1 + a_2 + (a_3 m_1 - a_2)\left(\dfrac{\theta_1 m_1 + \mu_1 - \lambda_1}{\mu_1 - \lambda_1}\right)^{\dfrac{\theta_2}{\theta_1}}}{a_4 m_1} + \frac{\lambda_2}{\theta_2}, \tag{5.3.1}$$

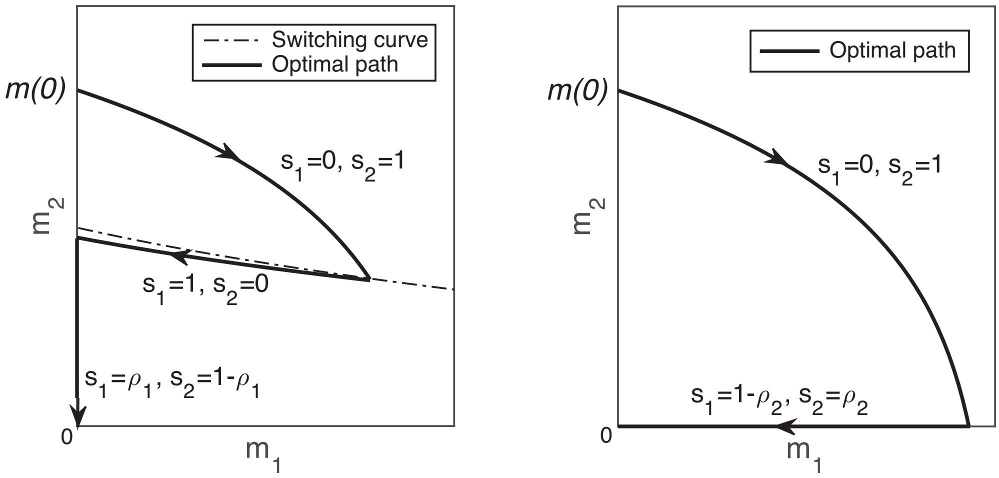
  den, eta

$$a_1 = \tilde{c}_2 \frac{\mu_2}{\theta_2}(1 - \rho); \quad a_2 = a_1 \frac{\mu_1}{\theta_1}(1 - \rho_1); \quad a_3 = -\left(\tilde{c}_2 \frac{\mu_2}{\theta_2} - \tilde{c}_1 \frac{\mu_1}{\theta_1}\right)(1 - \rho_1),$$

$$\text{and } a_4 = \left(\tilde{c}_2 \frac{\mu_2}{\theta_2} - \tilde{c}_1 \frac{\mu_1}{\theta_1}\right) \frac{\theta_2}{\mu_2}.$$

  Hau da, 2 klasea zerbitzatuko da $h(\cdot)$ trukatze-funtziora iritsi bitartean eta gero 1 klasea zerbitzatuko da $m_1 = 0$ den arte. Hortik aurrera, $m_1 = 0$ mantendu eta gainerako zerbitzua 2 klaseari emango zaio, ikusi 5.2a. Irudia.

- *$\tilde{c}_2\mu_2 \geq \tilde{c}_1\mu_1$ bada, orduan*



(a) Ibilbide optimoa $\tilde{c}_2\mu_2 \leq \tilde{c}_1\mu_1$ denean. (b) Ibilbide optimoa $\tilde{c}_2\mu_2 \geq \tilde{c}_1\mu_1$ denean.

Irudia 5.2: Politika optimoa $\dfrac{\tilde{c}_2\mu_2}{\theta_2} \geq \dfrac{\tilde{c}_1\mu_1}{\theta_1}$ den kasuan, eta ibilbide optimoa.

– $s^* = (0, 1)$, $m_2 > 0$ *denean,*

– $s^* = (1 - \rho_2, \rho_2)$, $m_2 = 0$ *denean.*

Hau da, 2 klasea zerbitzatu $m_2 = 0$ *den arte. Hortik aurrera,* $m_2 = 0$ *mantendu eta gainontzeko zerbitzua 1 klaseari eman, ikusi 5.2b. Irudia.*

Soluzioa $\tilde{c}_2\mu_2/\theta_2 \leq \tilde{c}_1\mu_1/\theta_1$ *kasuan baliokidea da indizeak aldatuta.*

5.1 Oharra (Klase kopurua arbitrarioa denean)**.** *$K = 2$ kasuan soluzioa aurkitzea konplikatua gertatu da, hori dela eta tesi honen autoreak ez dira K klase arbitrarioko kasuaren soluzio optimoa aurkitzen saiatu. Ordez, 5.4. Sekzioan heuristika bat garatu da $K = 2$ kasutik lortu den intuizioa aplikatuz.*

5.2 Oharra. *Ohartu, 5.1. Proposizioan definitu den $h(\cdot)$ trukatze-funtzioak absiza bertikala mozten duela*

$$h(0) = (1-\rho)\frac{\mu_2}{\theta_1\theta_2}\left(\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\dfrac{\tilde{c}_2\mu_2}{\theta_2} - \dfrac{\tilde{c}_1\mu_1}{\theta_1}}\right),$$

puntuan. Lan karga txikien kasuan egonik, azken hau positiboa dela ondoriozta daiteke, baldin eta soilik baldin

$$(\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2) \,/\left(\frac{\tilde{c}_2\mu_2}{\theta_2} - \frac{\tilde{c}_1\mu_1}{\theta_1}\right) \geq 0.$$

Azken proposizioaren frogapena 5.1. Lematik ondorioztatzen da. Lema hau aurkeztu aurretik soluzio optimoaren intuizioa azalduko da, zeina $\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ indize politikek karakterizatzen duten. Fluido kopurua nahiko txikia denean $\tilde{c}\mu$ indize poitika da optimoa. Hau hurrengo moduan azal daiteke. Ohartu kostuaren deribatua $\sum_{k=1}^{K} \tilde{c}_k \frac{\mathrm{d}m_k(t)}{\mathrm{d}t} = \sum_{k=1}^{K} \tilde{c}_k \left(\lambda_k - \mu_k s_k(t) - \theta_k m_k(t)\right)$ dela. $\tilde{c}\mu$ erregelak miopikoki minimizatzen du deribatua eta beraz optimoa da epe motzean. Jatorritik gertu hau da soluzio optimoak hartzen duen erabakia. Halere, epe luzean uzteen efektua ezin ahaz daiteke. Adibidez, $\tilde{c}_1\mu_1 > \tilde{c}_2\mu_2$ bada baina $\theta_1 >> \theta_2$, orduan erregela miopikoak 1 klaseari eskeiniko lioke lehentasuna. Honek $m_1(t)$ minimitzatzen du, zeinak kostuaren deribatuan influentzia negatiboa duen (cf. $\theta_1 m_1(t)$ gaia). Beraz, epe luzera ondo legoke 1 klaseko fluidoa kantitate handitan mantentzea, uzteen tasa oso handia baita.

5.1. Proposizioan erakutsi da jatorritik urruti dauden egoeretan, goian aipatu den efektua atzematen duen indizea $\tilde{c}\mu/\theta$ da. 5.2. Proposizioan ikusiko da $\tilde{c}\mu/\theta$ erregela optimoa dela $\rho > 1$ kasuan, lan karga handien kasuan.

$h(\cdot)$ trukatze-funtzioa, 5.1. Proposizioan definitu dena, $\tilde{c}_k\mu_k/\theta_k$ eta $\tilde{c}_k\mu_k$ erregelak zein egoeretan diren optimoak adierazten du. Hurrengoa ikas daiteke $h(\cdot)$-ren formulatik:

- 5.2. Oharrean ikus daitekeen bezela, $\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2$ eta $\dfrac{\tilde{c}_2\mu_2}{\theta_2} - \dfrac{\tilde{c}_1\mu_1}{\theta_1}$ espresioen arteko frakzioak $h(0)$ determinatzen du, eta beraz, trukatze-funtzioaren altuera. Hemendik ikus daiteke $\tilde{c}\mu$-en arteko diferentzia handitzen (txikitzen) den heinean $\tilde{c}\mu/\theta$-en arteko diferentziarekiko, trukatze-funtzioa ere handitu (txikitu) egiten da eta beraz, kontrol fluido optimoa $\tilde{c}\mu$ ($\tilde{c}\mu/\theta$) politikatik gerturatzen da.

- Lan karga 1-rantz hurbiltzen den heinean, *i.e.*, $\rho \uparrow 1$, $h(\cdot)$ trukatze-funtzioak $\overline{h}(\cdot)$-ra konbergitzen du non $\overline{h}(0) = 0$ den eta $\overline{h}(m_1) < 0$ edozein $m_1 > 0$ bada. Beraz, $\tilde{c}\mu/\theta$ indizea optimoa da eredu fluidoarentzat $\rho \uparrow 1$ denean. 5.3.2. Sekzioan ikusiko den bezela, $\tilde{c}\mu/\theta$ indize politika optimoa da lan karga handietarako ($\rho > 1$) ere, soluzio optimoaren jarraitasuna erakutsiz.

5.3 Oharra (Ilara klase-anitza *deadline*-ekin)**.** *$c_k = 0$ den kasuan edozein $k = 1, \ldots, K$ bada, eredua ilara klase-anitza deadline-ekin bihurtzen da: bezeroak beraien deadline-ak agortu aurretik zerbitzatu behar dira non deadline-ak θ_k parametroko banaketa esponentziala duten eta deadline-a amaitzen denean ez badute zerbitzua jaso sistema utziko dute δ_k kostua eraginez. Kasu partikular honetan $\tilde{c}\mu$ araua $\delta\mu\theta$ araua bilakatzen da eta $\tilde{c}\mu/\theta$ erregela $\delta\mu$ erregela.*

5.1. Proposizioa frogatzeko hurrengo lema beharrezkoa da. Honen frogapena 5.6.1. Eranskinean aurki daiteke. Lemak dio $\tilde{c}_k\mu_k$ dela lehentasuna erabakitzen duen indizea 1 klaseko eta 2 klaseko fluido kantitatea txikia denean.

5.1 Lema. *Izan bedi $K = 2$ eta $m(0) = (\varepsilon, \varepsilon)$ non $\varepsilon > 0$ eta txikia den. $\rho < 1$ bada eta*

$$\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2 \ (\textit{baliokideki } \tilde{c}_1\mu_1 \leq \tilde{c}_2\mu_2),$$

orduan 1 klaseak izango du lehentasuna (2 klaseak hurrenez hurren) jatorrira iritsi arte.

5.3.2 Lan karga handia denean politika optimoa bezero klase kopuru arbitrariorako

Sekzio honetan bezero klase kopuru arbitrario bat kontsideratuko da, *i.e.*, $K \geq 2$. P Problemaren analisiarekin amaitzeko $\rho > 1$ kasua falta da, zeinetarako helburua batez besteko kostua minimizatzea den (azken hau hertsiki positiboa da). Hurrengo proposizioak kontrok optimoa zehazten du eredu fluidoarentzat.

5.2 Proposizioa. *Izan bedi $\lambda_k, \mu_k, \theta_k, c_k$ eta δ_k finkoak edozein $k \in \{1, \ldots, K\}$, eta demagun klaseak hurrengo ordena jarraitzen dutela $\frac{\tilde{c}_1\mu_1}{\theta_1} \geq \frac{\tilde{c}_2\mu_2}{\theta_2} \geq \ldots \geq \frac{\tilde{c}_K\mu_K}{\theta_K}$. $\rho < 1$ bada, orduan P Problemarentzat $s^*(\cdot)$*

soluzio optimoa batez besteko kostua minimizatzen duena

$$s^*(t) = (\rho_1, \ldots, \rho_l, 1 - \sum_{i=1}^{l(t)} \rho_i, 0, \ldots, 0),$$

da, non $l(t) := \min\{k : m_{k+1}(t) > 0\}$ *den. Hau da, lehentasuna* $\tilde{c}\mu/\theta$ *indizeak erabakitzen du.*

Frogapena. Lehenik eta behin oreka puntu optimoa determinatu behar da. Oreka puntuak $0 = \lambda_k - \mu_k s_k - \theta_k m_k$, for all k betetzen du. Beraz, kontrol optimoa (orekan) oreka puntua minimizatzen duena

$$\arg\min_{s \in \mathcal{S}} \sum_{k=1}^{K} \tilde{c}_k m_k = \arg\min_{s \in \mathcal{S}} \sum_{k=1}^{K} \tilde{c}_k \frac{\lambda_k - \mu_k s_k}{\theta_k} = \arg\min_{s \in \mathcal{S}} \frac{\sum_{k=1}^{K} \tilde{c}_k \mu_k}{\theta_k} s_k,$$

definitzen du. Azken hau $\tilde{c}\mu/\theta$-ren araberako lehentasunak jarraituz minimizatzen da, hau da, oreka puntu optimoa $m^* = (0, \ldots, 0, \frac{\lambda_{j+1} - \mu_{j+1}(1 - \sum_{i=1}^{j} \rho_i)}{\theta_{j+1}}, \frac{\lambda_{j+2}}{\theta_{j+2}}, \ldots, \frac{\lambda_K}{\theta_K})$ da eta $s^* = (\rho_1, \ldots, \rho_j, 1 - \sum_{i=1}^{j} \rho_i, 0, \ldots, 0)$, non j-k $\sum_{i=1}^{j} \rho_i < 1$ betetzen duen, eta $\sum_{i=1}^{j+1} \rho_i \geq 1$.

Oraindik ikusteko dago $s^*(\cdot)$ kontrolpean dinamika fluidoak oreka egoerara konbergitzen duela. Hau hurrengo moduan ikus daiteke. Izan bedi $m^*(\cdot)$, $s^*(\cdot)$ kontrol aldagaiari dagokion ibilbidea. Izan bedi $w_j^*(t) := \sum_{k=1}^{j} m_k^*(t)/\mu_k$. $s^*(\cdot)$-en definizioz $\mathrm{d}w_j^*(t)/\mathrm{d}t = \sum_{k=1}^{j} \rho_k - 1 - \sum_{k=1}^{j} m_k^*(t)/\theta_k < -(1 - \sum_{k=1}^{j} \rho_k)$ $w_j^*(t) > 0$ denean lortzen da. Beraz, denbora finitu batean T prozesua 0-ra doa $w_j^*(T) = 0$, eta bertan geratzen da. Hortik aurrera, $j+1$ klaseari $1 - \sum_{k=1}^{j} \rho_k$ kapazitatea esleitzen zaio, halakorik bada. Beraz, zuzenean ikus daiteke sistemak m_{j+1}^* puntuan amaitzen duela, zeinak $0 = \lambda_{j+1} - \mu_{j+1}(1 - \sum_{k=1}^{j} \rho_k) - \theta_{j+1} m_{j+1}$ ebazten duen. $t > T$ denean $m_{j+1}^*(t) > 0$ da, eta $j+2, \ldots, K$ klaseek ez dute zerbitzurik jasotzen. Beraz, beraien dinaminak $\mathrm{d}m_i^*(t)/\mathrm{d}t = \lambda_i - \theta_i m_i^*(t)$, eta $m_i^*(t)$ balioak λ_i/θ_i-ra konbergitzen du edozein $i \in \{j+2, \ldots, K\}$. □

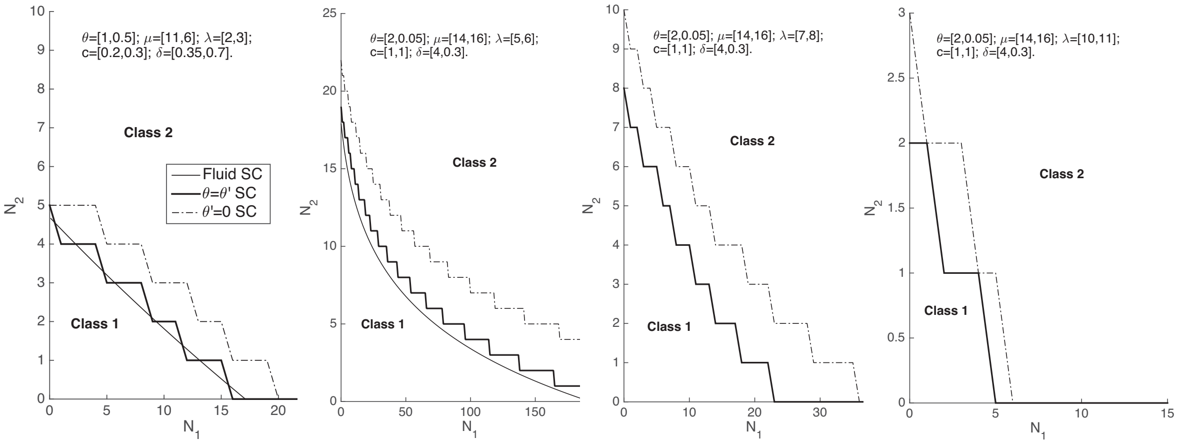
$\tilde{c}\mu/\theta$ erregela lehenago proposatua izan da Atar et al. [8, 9] artikuluetan, non *scheduling* optimoa aztertu den zerbitzari-anitzeko ilaretan. Arau hori hurbilketa fluido bat ebatziz lortu zen. Eredu fluidoa 5.2. Proposizioan aurkeztu denaren antzekoa da, baina baldintza gehiago ditu, $s_k \leq m_k$, zerbitzari-anitzen propietatea. Gainera, $\tilde{c}\mu/\theta$ indizea 3.3. Proposizioko Whittle indizearekin bat dator $c'_k = c_k$, $\theta_k = \theta'_k$ eta $\delta_k = \delta'_k$ kasurako.

5.3.3 Kontrol optimoa eredu estokastiko eta fluidoaren arteko alderatzea

Sekzio honetan eredu fluidoarentzat lortu den trukatze-funtzioa eta eredu estokastikoaren soluzio optimoa erkatuko dira $\theta = \theta'$ eta $\theta' = 0$ kasuetan. Soluzio optimoa *value iteration* algoritmoa erabiliz lortu da. 5.3. Irudian erkaketa hau egin da parametru multzo ezberdinetarako. Ohartu $\theta = \theta'$ ereduaren trukatze-funtzioa beti $\theta' = 0$ kasuko trukatze-funtzioaren azpitik dagoela. Izan ere, bezeroak sistema utzi badezakete zerbitzatuak izaten ari direnean, uzteen efektua handiagoa da.

5.3a. eta 5.3b. Irudietan $\rho < 1$ kasua kontsideratu da, $\rho = 1$ kasu kritikoa 5.3c Irudian eta $\rho > 1$ lan karga handiko kasua 5.3d. Irudian. Bestalde, 5.3b., 5.3c. eta 5.3d. Irudiak 5.5. Sekzioko 1 Adibideari dagozkie.

5.3a-5.3b Irudietan parametroek $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ and $\tilde{c}_1\mu_1/\theta_1 \leq \tilde{c}_2\mu_2/\theta_2$ betetzen dute, beraz, eredu fluidoaren soluzio optimoa trukatze-funtzio batek karakterizatzen du, non 2 klaseak duen lehentasuna



(a) $\rho = 0.6818$. (b) $\rho = 0.7321$. (c) $\rho = 1$. (d) $\rho = 1.4018$.

Irudia 5.3: Trukatze-funtzioa $\theta = \theta'$ eta $\theta' = 0$ kasuetan eta kontrol fluido eredua.

gainetik eta 1 klaseak azpitik. Ikusi daiteke eredu fluidoaren trukatze-funtzio optimoak eredu estokastikoaren trukatze-funtzio optimoa ongi hurbiltzen duela, konstante bat izan ezik zeina doiketa fluidoarekin desagertu egiten den.

5.3c. eta 5.3d. Irudietan eredu estokastikoaren soluzio optimoa trukatze-futnzio batek karakterizatzen du, zeinetan 2 klaseak duen lehentasuna funtzioaren gainean eta 1 klaseak azpian. Halere, eredu fluidoaren soluzio optimoak 2 klaseari ematen dio lehentasuna egoera-espazio guztian $\tilde{c}_1\mu_1/\theta_1 \leq \tilde{c}_2\mu_2/\theta_2$ eta $\rho > 1$ baitira. 5.3c. Irudian $\theta = \theta'$ kasuan soluzio optimopean batez besteko bezero kopurua $(\overline{m}_1, \overline{m}_2) = (0.7796, 4.1194)$ da zeina trukatze-funtzioaren azpitik dagoen. Beraz, politika estokastiko optimoak ia uneoro 1 klaseari emango dio lehentasuna. Eredu fluidoaren soluzioak ez du jatorrizko sistemaren funtsa atzematen, izan ere $h(\cdot)$ trukatze-funtzioa desagertu egiten da $\rho = 1$ denean. 5.5. Sekzioko 1 Adibidean soluzio fluidoaren optimoarekiko errore erlatiboa % 30-ekoa dela ikusiko da. $\rho > 1$ denean, kapitulu honetan proposatu den indizeak errendimendu ona erakusten du ikusi 5.5. Sekzioa. Izan ere, sistema trukatze funtzioaren gainetik bizi da. Adibidez, 5.3d. Irudian kontsideratu den parametroentzat, batez besteko bezero kopurua $(\overline{m}_1, \overline{m}_2) = (3.0088, 3.4849)$ da zeina trukatze-funtzioaren gainetik dagoen. Beraz, soluzio optimo estokastikoak lehentasuna 2 klaseari emango dio, zeinak fluidoaren soluzioarekin bat egiten duen.

5.4. Sekzioan eredu fluidoaren soluzioa nola erabil daitekeen jatorrizko eredu estokastikorako azalduko da. 5.5. Sekzioan numerikoki ebaluatuko da proposatutako heuristikaren errendimendua eredu estokastikoari aplikatzen zaionean. Errendimendua ona dela ikusiko da. Halere, ez da emaitza analitikorik lortu errore erlatiboari dagokionean. Literaturan, emaitza asintotikoki optimoak eredu fluidoentzat hainbat ilara ereduetako *scheduling* problematan lortu dira, ikusi adibidez [17, 43, 67, 68, 93]. Bereziki, ikusi da eredu fluidotik lortzen den soluzio optimoa eredu estokastikoari aplikatzen zaionean, doiketa-fluidoa jasan duen kostuak konbergitzen duela eredu fluidoaren kostura, azken hau behe borne bat bezala ikus daitekeelarik. Eredu partikular honetan eredu fluidoa hurbiketa soil bat bezela aurkeztu da, ez dago inongo bermerik eredu fluidoaren soluzioa eredu estokastikoan aplikatzean asintotikoki optimoa izango denik. Ohartu [8, 9] artikuluetan $\tilde{c}\mu/\theta$ erregela asintotikoki optimoa dela frogatu dela zerbitzari-anitzeko ilaran eta karga handietarako. Zerbitzari-anitzen kasua dela-eta, [8, 9] autoreek beste eremu limite bat azter dezakete:

iritsiera tasak eta zerbitzari kopurua doitzen direnean eta zerbitzu tasak finko mantentzen direnean μ tasarekin. Espero da lan karga handietarako antzeko frogapen bat egin ahal izango litzatekeela kapitulu honetako ereduan.

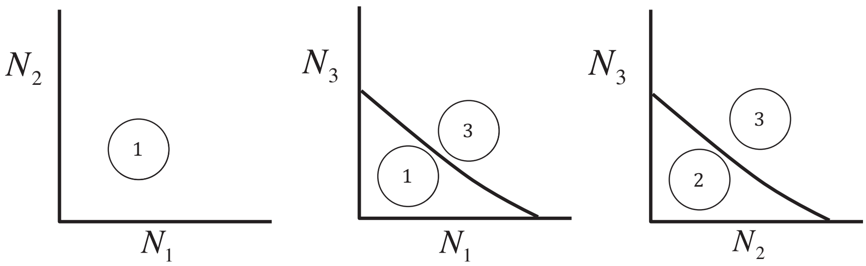
5.4 Bezero klase kopuru arbitrarioa

Sekzio honetan uzteak gerta daitezkeen optimizazio eredu estokastikoarentzat heuristika bat proposatuko da. Heuristika hau eredu fluidoa ebaztean lortu den intuizioan oinarritzen da.

Lehenik eta behin lan-karga handien eremua kontsideratuko da. Kasu honetan, eredu fluidoaren soluzio optimoa lehentasuna $\tilde{c}\mu/\theta$ indizearen arabera ematea da. 5.5. Sekzioan politika honen ebaluatzea egingo da eredu estokastikoari aplikatuz.

Orain lan-karga txikietako kasua kontsideratuko da. Gogoratu 5.1. Proposizioan soluzio optimoak egitura berezi bat duela ikusi dela bi bezero klase diren kasuan: jatorritik gertu $\tilde{c}\mu$ indize politika da optimoa, eta klase bateko fluido kopurua nahiko handia denean $\tilde{c}\mu/\theta$ indize politika da optimoa. Eredu estokastikoa numerikoki ebatziz, soluzio optimoak egitura bera duela ikus daiteke, ikusi 5.3.3. Sekzioa eta 5.5. Irudia (ezkerraldean). Kapitulu honetan erabili den hurbilketa ideia bera jarraituz heuristika bat garatzea da, hau da, jatorritik gertu lehentasuna $\tilde{c}\mu$ politikaren arabera ezartzea eta bezeroen kopurua nahiko handia denean $\tilde{c}\mu/\theta$ politika jarraitzea. Ez dago argi zein den aukera egokia trukatze-funtziotik gertu $\tilde{c}\mu$ ala $\tilde{c}\mu/\theta$.

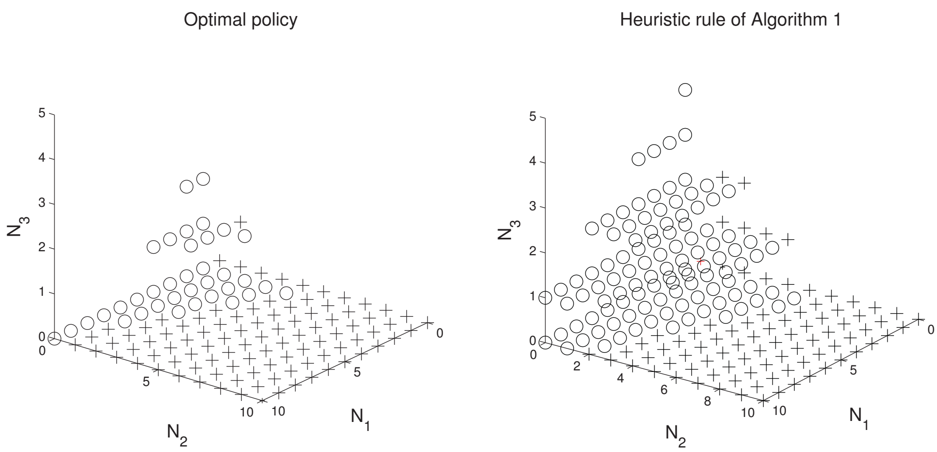
Hurrengo heuristika proposatuko da, zeina bi klaseko eredu fluidoaren analisian oinarritzen den: K klaseko ilara orokor batentzat klase guztiak binaka-binaka konparatuko dira eta eurei dagokien trukatze-funtzioak kalkulatuko dira, ikusi adibidez 5.4. Irudia zeinarentzat $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2 \geq \tilde{c}_3\mu_3$ eta $\tilde{c}_3\mu_3/\theta_3 \geq \tilde{c}_1\mu_1/\theta_1 \geq \tilde{c}_2\mu_2/\theta_2$. Orduan, $(N_1, N_2, \ldots, N_k)$ egoerak edozein bi (N_i, N_j) gai hartuz, trukatze-funtzioaren azpitik daudela betetzen badu, orduan lehentasuna $\tilde{c}\mu$ handiena duen klaseari emango zaio. Baina, existitzen bada (N_i, N_j) bikoterik zeina eurei dagokien trukatze-funtzioaren gainetik dagoen, orduan lehentasuna $\tilde{c}\mu/\theta$ indizearen araberakoa izango da. Adibidez, 5.4. Irudiko parametroak hartuz gero, lehentasuna 1 klaseari dagokio (N_2, N_3) eta (N_1, N_3) bikote bakoitzari dagokion trukatze-funtzioaren azpitik badaude eta kontrako kasuan 3 klaseari. Klase bateko ilara hutsa dagoenean sistemak gainontzeko $K-1$ ilarak bakarrik hartuko ditu kontuan. Ulergarritasuna hobetzearren hurrengo 1.*pseudo*-algoritmoan heuristika aurkeztuko da



Irudia 5.4: $K = 3$ deneko heuristikaren adibide bat $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2 \geq \tilde{c}_3\mu_3$ eta $\tilde{c}_3\mu_3/\theta_3 \geq \tilde{c}_1\mu_1/\theta_1 \geq \tilde{c}_2\mu_2/\theta_2$ direnean.

Algorithm 1 Algoritmoa K arbitrario baterako heuristika kalkulatzeko

Demagun r ilara ez hutsak direla
Izan bedi N_i $i \in \{1, \ldots, r\}$ klaseari dagokion egoera.
Kalkulatu $\tilde{c}_i\mu_i$ eta $\tilde{c}_i\mu_i/\theta_i$ indizeak edozein $i \in \{1, \ldots, r\}$.
i eta j bikote bat emanik, non $\tilde{c}_i\mu_i/\theta_i \geq \tilde{c}_j\mu_j/\theta_j$ den, kalkulatu (5.3.1). Ekuazioan definitutako h_{ij} trukatze-funtzioa
if edozein $i, j, N_i \leq h_{ij}(N_j)$ **then**
  Lehentasuna eman $\tilde{c}\mu$ handiena duen klaseari
else
  Lehentasuna eman $\tilde{c}\mu/\theta$ handiena duen klaseari
end if



Irudia 5.5: Hiru klaseko adibide baterako politika optimoa eta heuristika $\theta' = 0$ kasuan. Zirkuluak dauden eremuan 1 klasea da zerbitzatua, batu ikurrreko eremuan 2 klaseak du lehentasuna, eta ikurrik gabeko eremuan 3 klaseak du lehentasuna.

Definitu berri den heuristika soluzio optimoarekin (numerikoki lortu dena *value iteration* algoritmoa erabiliz) konparatzeko adibide bat proposatu da, zeinarentzat $K = 3$ den. Kontsideratu $\mu = [10, 10, 9]; \theta = [1, 0.5, 0.25]; c = [1.7, 1.7, 1.7]; \delta = [2, 2, 4]; \lambda = [2, 2, 1]$ parametroak. Beraz, $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2 \geq \tilde{c}_3\mu_3$ eta $\tilde{c}_1\mu_1/\theta_1 \leq \tilde{c}_2\mu_2/\theta_2 \leq \tilde{c}_3\mu_3/\theta_3$. Kapitulu honetako heuristika kontuan hartuz, 1 klaseak izango du lehentasuna hiru klaseak jatorritik gertu daudenean, hau da, $\tilde{c}\mu$ indize politikak erabakiko du lehentasuna, eta 3 klaseak izango du lehentasuna beste kasuan, hau da, $\tilde{c}\mu/\theta$ indize politikak erabakiko du lehentasuna. 2 klaseak izango du lehentasuna hurrengo bi kasuetan: (i) 1 klasea hutsa denean eta (N_2, N_3) jatorritik nahiko gertu dagoenean ($\tilde{c}\mu$ indize politikaren arabera), eta (ii) 3 klasea hutsa denean eta (N_1, N_2) jatorritik nahiko urruti dagoenean ($\tilde{c}\mu/\theta$ indize politikaren arabera).

5.5. Irudian (ezkerraldean) *value iteration* bidez kalkulatu den soluzio optimoa irudikatu da eta heuristikak emandako soluzioa (eskuinaldean). Ikus daiteke heuristikak kualitatiboki antzeko emaitza eskeintzen duela. 5.5.2. Sekzioan konparaketa numeriko bat egingo da errendimendua aztertzeko.

5.5 Emaitza numerikoak

Sekzio honetan numerikoki ebaluatu da 1. Algoritmoan deskribatutako heuristikaren errendimendua. Heuristikaren errendimendua soluzio optimoarekin erkatu da. Azken hau 1.3.3. Sekzioan aurkeztu den *value iteration* algoritmoa erabiliz lortu da. Hurrengo indize politikak ere ebaluatu dira:

- $\tilde{c}\mu/\theta$ erregela. Politika hau [8, 9] artikuluetan aurkeztu da zeinetan asintotikoki optimoa dela frogatu den zerbitzari-anitzeko sistemetan eta lan karga handietarako. 5.2. Proposizioan ikusi da kapitulo honetako eredu fluidoarentzat lan karga handietarako ere optimoa dela.
- $\tilde{c}\mu/\theta - c$ erregela. Politika hau [15] artikuluan garatu da $\theta' = 0$ kasurako (iritsierarik gabeko sistema batean) eta zerbitzuan dagoen bezeroak ere kostu bat eragiten duen kasuan.
- $\tilde{c}\mu$ erregela. Hau erregela miopikoa da zeinak berehalako kostua minimizaten duen. Politika hau $c\mu$ ([33]) erregelaren analogotzak ikusi daiteke uzterik gabeko sistementzat.

$\tilde{c}\mu/\theta$ eta $\tilde{c}\mu/\theta - c$ indize politikak bat datoz 3.3. Proposizioan lortu diren Whittle indize politikekin. $\tilde{c}\mu$ politika 3.7. Proposizioko Whittle indize politikarekin dator bat mantentze-kostu linearren kasuan.

Emaitzak azaldu aurretik, analisi honetatik atera diren ondorioak aurkeztuko dira:

- $\theta = \theta'$ eta $\theta' = 0$ ereduak kualitatiboki oso antzekoak dira.
- $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu/\theta - c$ indize politikak errendimendu ona erakusten dute lan karga handietarako.
- 1. Algoritmoan aurkeztu den heuristikak errendimendu ona erakusten du lan karga ezberdinetarako.

Zuzentasunez, aipatu behar da, heuristikak, $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu$ indize politikek baino errendimendu hobea erakusten duen arren, heuristika fluidoa inplementatzea zailagoa da indize politikak erabiltzea baino.

Orain kapitulu honetan aztertu diren egoerak aurkeztuko dira. 5.5.1. Sekzioan $K = 2$ kasua kontsideratu da eta 5.5.2. Sekzioan $K = 3$ kasua.

5.5.1 Errendimenduaren ebaluatzea bi bezero klaserentzat

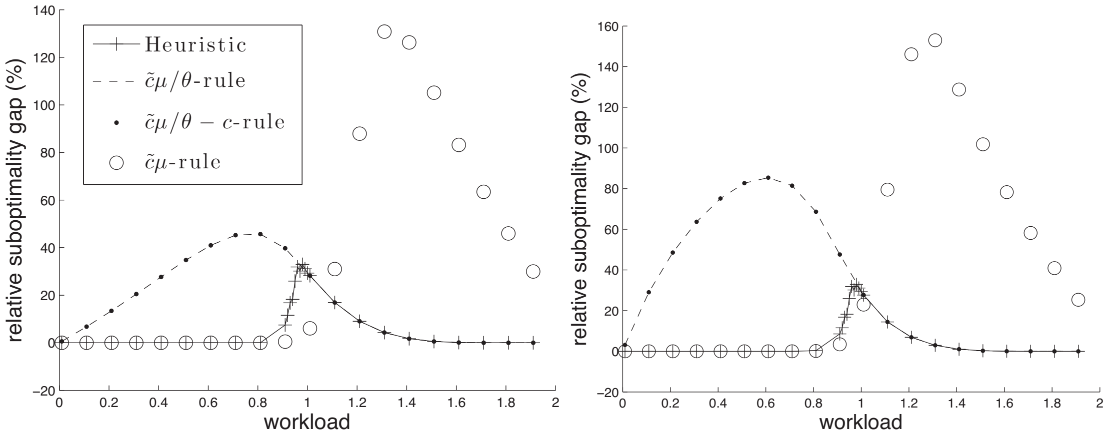
Bi ereduak aztertuko dira, $\theta = \theta'$ eta $\theta' = 0$, eta errore erlatiboa kalkulatuko da goian deskribatu diren politikentzat. 1. eta 2. Adibideetan c, δ, μ, θ eta $\rho_1 = \rho_2$ parametroak finkatu dira eta ρ-k balioak aldatzen direla kontsideratu. 3. Adibidean ρ finkoa da eta θ_1 aldatzen da.

1 Abidibea: Adibide honetan finkatu $\theta = [2, 0.05]; \mu = [14, 16]; c = [1, 1]; \delta = [4, 0.3]$, non $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ eta $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$. Adibide honei dagozkien emaitzak 5.6a. Irudian aurki daitezke $\theta = \theta'$ kasurako eta 5.6b. Irudian $\theta' = 0$ kasurako.

Lan karga txikietan heuristikak eta $\tilde{c}\mu$ indize politikak errendimendu optimoa erakusten dute, ordez $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu/\theta - c$ indize erregelek errendimendu kaxkarra dute. 5.3b. Irudian $\rho = 0.73$ kasuari dagokion trukatze-funtzioa irudikatu da. Batez besteko bezero kopuruak ($\theta = \theta'$ kasuan eta $\rho = 0.73$ denean) $(\overline{N}_1, \overline{N}_2) = (0.4047, 1.5092)$ balioa hartzen du, zeina $\theta = \theta'$ ereduari dagokion trukatze-funtziotik eta eredu fluidoaren trukatze-funtziotik urruti dagoen. Honek erakusten du heuristika eta $\tilde{c}\mu$ politika optimoak izatearen arrazoia (eurek atzematen baitute kontrol optimoa trukatze-funtzioaren azpitik).

Lan karga handietan, $\tilde{c}\mu$ indize politikak errore handia erakusten du eta $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu/\theta - c$ indize politikak ia optimoak dira. Azken hau itxarotako emaitza da 5.3.3. Sekzioan eztabaidatu den legez.

Ohartu lan karga $\rho = 1$ denean errendimenduaren errore erlatiboa % 30-ekoa da. Kapitulo honetan garatu den heuristikak 2 klaseari emango dio lehentasuna $\rho > 1$ denean eta trukatze-funtzio bat erakutsiko



(a) Errore erlatiboa $\theta = \theta'$ kasuan (b) Errore erlatiboa $\theta' = 0$ kasuan

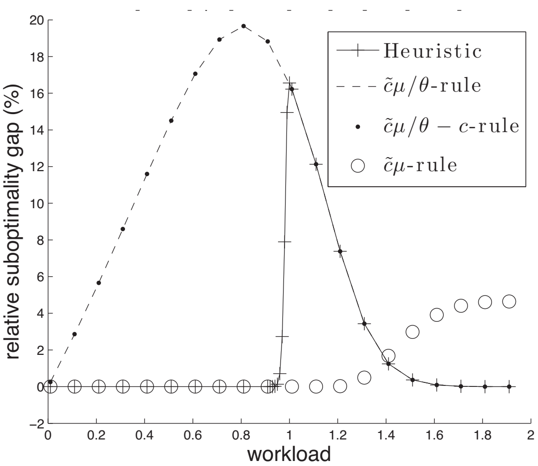
Irudia 5.6: 1. Adibideari dagokion politiken errendimenduen konparaketa.

du jatorritik oso gertu dagoena $\rho = 1 - \epsilon$ denean. 5.3c. Irudian, adibide honi dagokiona, ikus daiteke eredu estokastikoak trukatze-funtzio bat duela $\rho = 1$ denean. Beraz, trukatze-funtzioaren azpitik dagoen egoera batean 1 klaseak izango du lehentasuna. Prozesua $\rho = 1$ denean batez beste trukatze-funtziotik gertu bizi da, beraz, kapitulo honetan garatu den politika optimotik urruti egon daiteke, 5.3.3. Sekzioan eztabaidatu den legez.

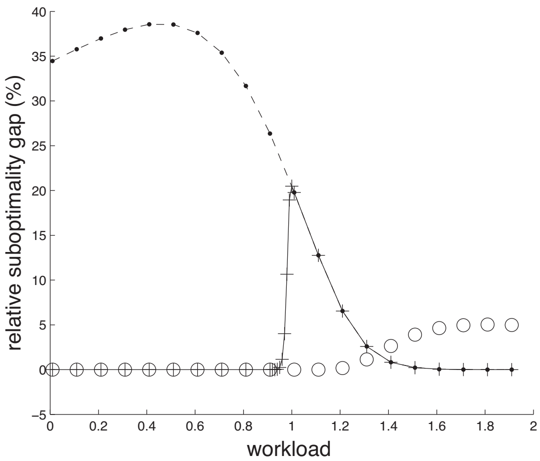
2. Adibidea: Bigarren adibide honetan $\theta = [1, 0.5]; \mu = [15, 25]; c = [0, 0]; \delta = [5, 3.2]$, dira eta beraz, $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ eta $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$. 5.3. Oharrean azaldu den legez, $c_1 = c_2 = 0$ eginez gero ereduak beste interpretazio bat du: bezeroak euren epemuga amaitzen denean sistema uzten dute zerbitzua jasotzen amaitu baino lehen. Kasu honetan $\tilde{c}\mu/\theta$, $\tilde{c}\mu/\theta - c$ eta $\tilde{c}\mu$ indize politikak $\delta\mu$, $\delta\mu$, eta $\delta\theta\mu$ indize politiken baliokideak dira, hurrenez hurren. Ohartu $\delta\mu$ indizeak ez duela errendimendu ona erakusten lan karga txikietan baina ia optimoa da lan karga handietarako. $\delta\theta\mu$ indize politikarekin kontrakoa gertatzen da. Kapitulo honetan garatu den heuristika optimoa da lan karga txikietarako $\delta\mu$ indizea bezain ona, ikusi 5.7a. Irudia $\theta = \theta'$ kasurako eta 5.7b. Irudia $\theta' = 0$ kasuan.

5.8. Irudian eredu estokastikoaren trukatze-funtzioa irudikatu da $\theta = \theta'$ eta $\theta' = 0$ kasuetarako (*value iteration* algoritmoa erabiliz lortu dena), eta eredu fluidoaren trukatze-funtzioa $h(\cdot)$. 5.8a. Irudia $\rho = 0.8867$ kasuari dagokio. Kasu horretan, batez besteko bezero kopurua $(\overline{N}_1, \overline{N}_2) = (0.6859, 2.6963)$ da, zeina trukatze-funtzioen azpitik dagoen. Beraz, honek erakusten du heuristika fluidoa eta $\tilde{c}\mu$ indize politikak zergatik diren ia optimoak. Bestalde, $\rho = 1$ denean, 5.8b. Irudian, eredu fluidoarean kontrol optimoa 2 klasea zerbitzatzea da, eta beraz ez dago trukatze-funtziorik. Eredu estokastikoaren kontrol optimopean, batez besteko bezero kopurua $(\overline{N}_1, \overline{N}_2) = (0.763, 4.2703)$ da. Hau trukatze-funtzioaren azpitik dagoen egoera bat da. Beraz, uneoro lehentasuna 1 klaseari ematen zaio. Honek azaltzen du heuristika fluidoak duen errorea % 16-koa izatea. Halere, sistemaren lan karga handitzen den heinean ($\rho > 1$) prozesua trukatze-funtzio estokastikoaren gainetik biziko da. Ikusi adibidez 5.8c. Irudia $\rho = 1.52$ denean, zeinentzat bezero kopurua politika optimopean $(\overline{N}_1, \overline{N}_2) = (6.8054, 3.7244)$ den. Honek azaltzen du heuristika fluidoa, 2 klaseari lehentasuna ematea, ia optimoa izatea.

5.4 Oharra (Lan-karga 1-etik gertu dagoeneko tontorra)**.** *Ohartu 1. eta 2. adibideetan lan-karga 1-etik gertu dagoenean tontor bat agertzen dela errore erlatiboan heuristika fluidorako. Hau hurrengo moduan*

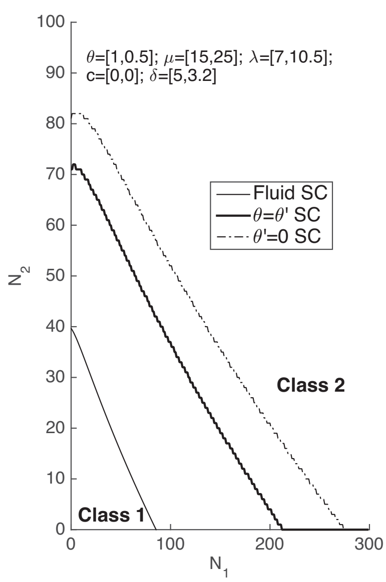


(a) Errore erlatiboa $\theta = \theta'$ kasuan

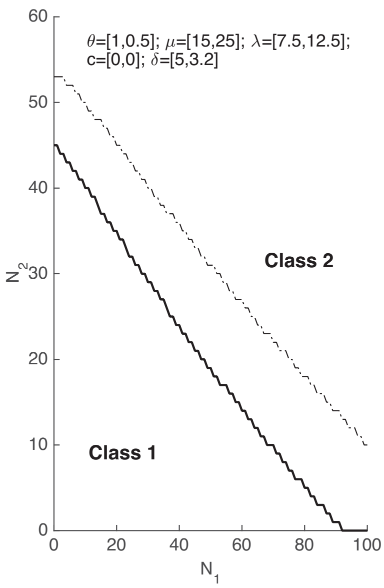


(b) Errore erlatiboa $\theta' = 0$ kasuan

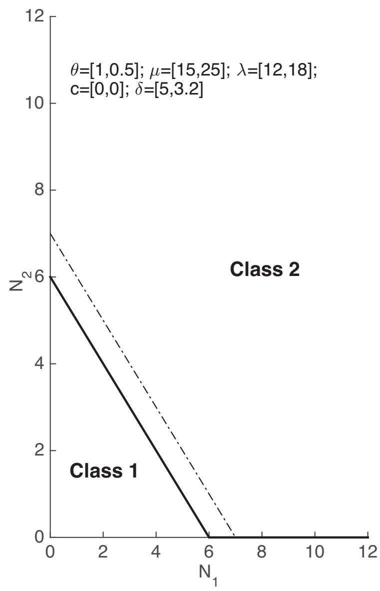
Irudia 5.7: 2. Adibideari dagokion politiken errendimenduen konparaketa.



(a) $\rho = 0.8867$.

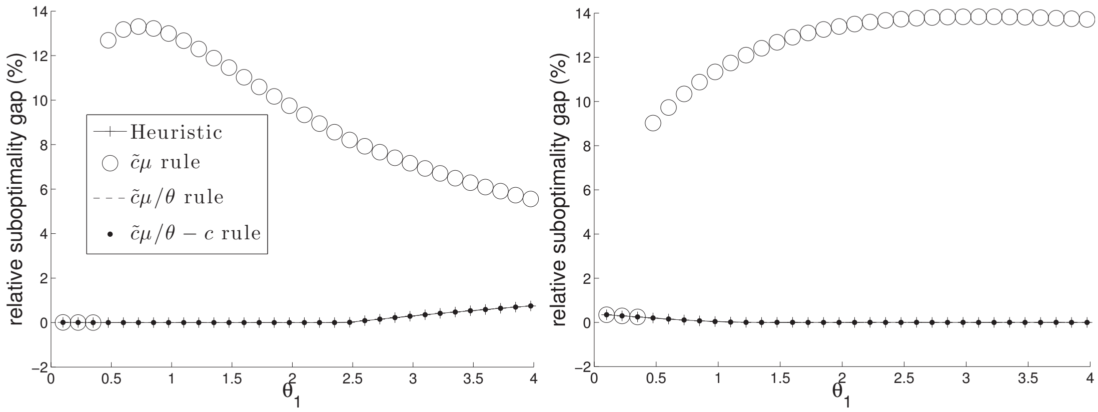


(b) $\rho = 1$.



(c) $\rho = 1.52$.

Irudia 5.8: 2. Adibideko trukatze-funtzioen konparaketa.



(a) Errore erlatiboa $\theta = \theta'$ kasuan

(b) Errore erlatiboa $\theta' = 0$ kasuan

Irudia 5.9: 3. Adibideari dagokion politiken errendimenduaren konparaketa, lan karga $\rho = 0.7$ denean.

azal daiteke. 5.1. Proposizioaren frogapenean, ikusi 5.6.2. Eranskina, trukatze-funtzio bat ikusi daiteke

$$h(0) = (1 - \rho_1 - \rho_2)\frac{\mu_2}{\theta_1\theta_2}\left(\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\frac{\tilde{c}_2\mu_2}{\theta_2} - \frac{\tilde{c}_1\mu_1}{\theta_1}}\right) > 0,$$

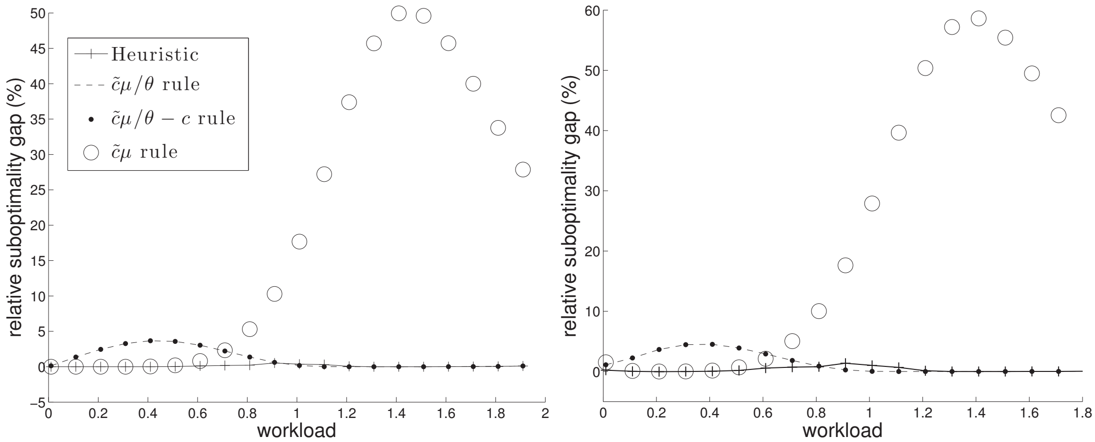
den kasuetan. Beraz, $1-\rho_1-\rho_2 \to 0$ *den heinean trukatze-funtzioa desagertu egiten da, hau da, heuristika* $\tilde{c}\mu/\theta$ *politikaren baliokidea da. Halere,* $\rho = 1$ *denean eredu estokastikoaren kontrol optimoa* $\tilde{c}\mu$ *politika jarraitzea da eta hutsa ez den egoera-espazioa zati batean, ikusi adibidez 5.8b. Irudia.*

3. Adibidea: Hurrengo parametroak kontsideratu dira: $\theta_2 = 0.1; \mu = [8, 8]; \lambda = [2.8, 2.8]; c = [1, 1]; \delta = [0.5, 2]$, eta demagun θ_1 aldatzen dela. Beraz, $\rho = 0.7$, *i.e.*, lan-karga txikien kasua da. Emaitzak 5.9a. Irudian aurki daitezke $\theta = \theta'$ kasurako eta 5.9b. Irudian $\theta' = 0$ kasurako.

$\theta_1 \in [0, 0.4]$ denean, $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$ da eta $\tilde{c}_2\mu_2 \geq \tilde{c}_1\mu_1$, zeinetarako heuristika fluidoak 2 klaseari ematen dion lehentasuna, baita indize politikek ere. Bestalde, $\theta_1 \in (0.4, 4]$ denean, $\tilde{c}_2\mu_2/\theta_2 \geq \tilde{c}_1\mu_1/\theta_1$ eta $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ eta trukatze-funtzio bat agertzen da heuristikarekin. Trukatze-funtziorik agertzen ez den kasuan ($\theta_1 \in [0, 0.4]$), indize politika guztiak optimoak dira, baina trukatze funtzio bat agertzen denean heuristikarentzat $\tilde{c}\mu$ erregelak errore positibo bat ematen du. $\tilde{c}\mu$ indize politikak errendimendu kaxkarra izatearen arrazoia $\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\tilde{c}_2\mu_2/\theta_2 - \tilde{c}_1\mu_1/\theta_1}$ frakzioa txikia bilakatzen den heinean, trukatze-funtzioa zerorantz hurbiltzen dela da eta beraz $\tilde{c}\mu/\theta$ optimo bihurtzen dela.

5.5.2 Errendimenduaren analisia bezero klase kopuru arbitrariorako

5.4. Sekzioan azaldu den heuristikaren errendimendua analizatuko da hemen. Kasu honetan 5.4. Sekzioan aztertu den aibide berdina aztertuko da hurrengo parametroekin: $\mu = [10, 10, 9]; \theta = [1, 0.5, 0.25]; c = [1.7, 1.7, 1.7]; \delta = [2, 2, 4]$. Izan bedi $\lambda_i = \lambda\beta_i$, $i = 1, 2, 3$, non λ iritsiera totala den eta β_i i klasearen bezero frakzioa. β_i, $i = 1, 2, 3$ $\rho_1 = \rho_2 = \rho_3$ bete daitezen aukeratuko dira. λ-ren balioa aldatuz lan karga alda daiteke eta errore erlatiboa kalkulatu politika ezberdinpean. Emaitzak 5.10. Irudian aurki daitezke. Lan-karga txikia denean heuristika fluidoa zein $\tilde{c}\mu$ indize politika optimoak dira, baina ρ handitzen den



(a) Errore erlatiboa $\theta = \theta'$ kasuan (b) Errore relatiboa $\theta' = 0$ kasuan

Irudia 5.10: Politika ezberdinen errendimenduen konparaketa $K = 3$ kasuan.

heinean $\tilde{c}\mu$ indize politikaren errendimendu kaxkartzen da. Bestalde, $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu/\theta - c$ indize politikak optimo bihurtzen dira lan karga 1 baino handiagoa denean. Heuristika fluidoak errore erlatibo txikia du.

5.6 Eranskina

5.6.1 5.1. Lemaren frogapena

5.1. Lemak jatorritik gertu dauden egoeretan zein klase zerbitzatzea den optimoa dio. Kostu funtzioa $(m_1(0), m_2(0)) = (\varepsilon, \varepsilon)$ hasierako puntutik hasita kalkulatu da 1 klaseari lehentasuna ematen zaionean. Berdina egin da lehentasuna 2 klaseari emango balitzaio kalkulua eginez. Bi kostu funtzioak konparatu dira eta 1 klaseari lehentasuna emateko behar det baldintza lortu da. Nahikoa da horretarako goian aurkeztu diren politikak konparatzea izan ere, kontrola lineala denez onar dezakegu optimoa izango dela prioritate osoa beti klase bati ala besteari ematea, jatorritik nahiko gertu hasiz gero, ikusi [34].

Lehenik eta behin 1 klaseari lehentasun osoa ematen dion kontrola kontsideratu da. 1 klaseak zero ikutzen duenean, ρ_1-eko edukiera eskeintzen zaio 1 klaseari eta $1 - \rho_1$-ekoa 2 klaseari, $(0, 0)$ oreka egoera iritsi arte. Politika honen pean kostua, $(m_1(0), m_2(0)) = (\varepsilon, \varepsilon)$ puntuan hasiz gero, hurrengoa da:

$$C_1(t, m) := \int_0^T \tilde{c}_1 m_1(t) + \tilde{c}_2 m_2(t) \, \mathrm{d}t.$$

$m_1(t)$ eta $m_2(t)$ traiektoriak kalkulatu ahal izateko, denbora bi tartetan banatuko da, $[0, t_1]$ eta $[t_1, t_2]$, non t_1 1 klaseak zero ikutzen duen unea den eta t_2 2 klaseak zero ikutzen duen unea. Hainbat kalkulu egin ostean, hurrengoa lor daiteke $[0, t_1]$ denbora tartean:

$$\begin{cases} m_1(t) = \left(\varepsilon + \dfrac{\mu_1 - \lambda_1}{\theta_1}\right) \mathrm{e}^{-\theta_1 t} + \dfrac{\lambda_1 - \mu_1}{\theta_1} = -\theta_1 t \left(\varepsilon + \dfrac{\mu_1 - \lambda_1}{\theta_1}\right) + \varepsilon + o(\varepsilon), & t \in [0, t_1], \\ m_2(t) = \left(\varepsilon - \dfrac{\lambda_2}{\theta_2}\right) \mathrm{e}^{-\theta_2 t} + \dfrac{\lambda_2}{\theta_2} = \theta_2 t \left(\dfrac{\lambda_2}{\theta_2} - \varepsilon\right) + \varepsilon + o(\varepsilon), & t \in [0, t_1]. \end{cases}$$

Hemen $t_2 \leq \frac{m_1(0)/\mu_1+m_2(0)/\mu_2}{1-\rho} = O(\varepsilon)$ [2] dela erabili da. Beraz, $\mathrm{e}^{-\theta_1 t} = -\theta_1 t + 1 + o(\varepsilon)$ edozein $t \leq t_2$ bada. (Hemen, $o(\varepsilon) = g(\varepsilon)$ non $g(\cdot)$ funtzioak $\lim_{\varepsilon \to 0} g(\varepsilon)/\varepsilon = 0$ betetzen duen). ε nahiko txikia denez $m_2(t) > 0$ edozein $t < t_1$ bada. t_1 denbora 1 klasea husten den denbora denez, hurrengoa lor daiteke

$$t_1 = \frac{\varepsilon}{\theta_1\left(\varepsilon + \dfrac{\mu_1 - \lambda_1}{\theta_1}\right)} = \frac{\varepsilon}{\mu_1 - \lambda_1} + o(\varepsilon),$$

orduan

$$m_2(t_1) = \frac{\lambda_2}{\theta_2}\frac{\varepsilon\theta_2}{\mu_1 - \lambda_1} + \varepsilon + o(\varepsilon). \tag{5.6.1}$$

Gogoratu t_2 2 klasea husten den denbora dela. $[t_1, t_2]$ denbora tartean 1 klaseak ρ_1 zerbitzua jasotzen du eta 2 klaseak $1 - \rho_1$. Beraz, hainbat kalkuluren ostean

$$\begin{cases} m_1(t) = 0, & t \in [t_1, t_2], \\ m_2(t) = A_2' \mathrm{e}^{-\theta_2 t} + \dfrac{\lambda_2 - \mu_2(1-\rho_1)}{\theta_2} = A_2'\left(-\theta_2 t + 1\right) + \dfrac{\lambda_2 - \mu_2(1-\rho_1)}{\theta_2} + o(\varepsilon), & t \in [t_1, t_2], \end{cases}$$

non A_2' integrazio konstantea den. Hemen $t = O(\varepsilon)$ erabili da, beraz, $\mathrm{e}^{-\theta_2 t} = -\theta_2 t + 1 + o(\varepsilon)$. Gainera, (5.6.1) Ekuaziotik

$$A_2' = \frac{-\dfrac{\lambda_2}{\theta_2}\left(\dfrac{-\varepsilon\theta_2}{\mu_1 - \lambda_1}\right) + \varepsilon + \dfrac{\mu_2(1-\rho_1) - \lambda_2}{\theta_2}}{1 - \dfrac{\theta_2 \varepsilon}{\mu_1 - \lambda_1}} + o(\varepsilon),$$

lortzen da. Beraz,

$$\begin{aligned} m_2(t) &= \frac{(\lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1))\varepsilon + (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)}{-\theta_2^2\varepsilon + (\mu_1 - \lambda_1)\theta_2}(-\theta_2 t + 1) + \frac{\lambda_2 - \mu_2(1-\rho_1)}{\theta_2} + o(\varepsilon), \\ &= \frac{(\lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1))\varepsilon + (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)}{-\theta_2^2\varepsilon + (\mu_1 - \lambda_1)\theta_2}(-\theta_2 t) \\ &\quad + \frac{(\mu_1 - \lambda_1 + \mu_2(1-\rho_1))\varepsilon}{-\theta_2\varepsilon + \mu_1 - \lambda_1} + o(\varepsilon), \qquad t \in [t_1, t_2], \end{aligned}$$

eta $m_2(t_2) = 0$ denez

$$\begin{aligned} t_2 &= \frac{(\mu_1 - \lambda_1 + \mu_2(1-\rho_1))\varepsilon}{-\theta_2^2\varepsilon^2 + (\lambda_2\theta_2 + \theta_2(\mu_1 - \lambda_1))\varepsilon + (\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)} + o(\varepsilon) \\ &= \frac{(\mu_1 - \lambda_1 + \mu_2(1-\rho_1))\varepsilon}{(\mu_2(1-\rho_1) - \lambda_2)(\mu_1 - \lambda_1)} + o(\varepsilon), \end{aligned}$$

[2] $w(t) := m_1(t)/\mu_1 + m_2(t)/\mu_2$ denez norabide negatiboa du zeina $\rho - 1$ baino txikiagoa den, ikusi 5.3. Sekzioko ohin-oharra

lortzen da. Orain kostua kalkula daiteke:

$$\begin{aligned}
C_1(t,(\varepsilon,\varepsilon)) &= \int_0^{t_1} \tilde{c}_1 \left(\left(\varepsilon + \frac{\mu_1-\lambda_1}{\theta_1} \right)(-\theta_1 t) + \varepsilon \right) + \tilde{c}_2 \left(\left(\varepsilon - \frac{\lambda_2}{\theta_2} \right)(-\theta_2 t) + \varepsilon \right) \mathrm{d}t \\
&\quad + \int_{t_1}^{t_2} \tilde{c}_2 \left(\frac{(\lambda_2\theta_2 + \theta_2(\mu_1-\lambda_1))\varepsilon + (\mu_2(1-\rho_1)-\lambda_2)(\mu_1-\lambda_1)}{-\theta_2^2\varepsilon + (\mu_1-\lambda_1)\theta_2}(-\theta_2 t) \right) \mathrm{d}t \\
&\quad + \int_{t_1}^{t_2} \tilde{c}_2 \left(\frac{(\mu_1-\lambda_1+\mu_2(1-\rho_1))\varepsilon}{-\theta_2\varepsilon + \mu_1 - \lambda_1} \right) \mathrm{d}t + o(\varepsilon^2) \\
&= \varepsilon^2 \left(\frac{\tilde{c}_1}{2(\mu_1-\lambda_1)} + \tilde{c}_2 \frac{2(\mu_1-\lambda_1)+\lambda_2}{2(\mu_1-\lambda_1)^2} \right) \\
&\quad - \tilde{c}_2 \frac{(\lambda_2\theta_2 + \theta_2(\mu_1-\lambda_1))\varepsilon + (\mu_2(1-\rho_1)-\lambda_2)(\mu_1-\lambda_1)}{2(-\theta_2\varepsilon + (\mu_1-\lambda_1))} \left((t_2)^2 - (t_1)^2 \right) \\
&\quad + \tilde{c}_2 \left(\frac{(\mu_1-\lambda_1+\mu_2(1-\rho_1))\varepsilon}{-\theta_2\varepsilon + \mu_1 - \lambda_1} \right)(t_2 - t_1) + o(\varepsilon^2),
\end{aligned}$$

non

$$(t_2)^2 - (t_1)^2 = \frac{(b_1\varepsilon^2 + b_2\varepsilon)^2}{(-b_1^2\varepsilon^2 + b_3\varepsilon + b_4)^2} - \varepsilon^2 b_5^2 + o(\varepsilon^2) = \frac{b_2^2\varepsilon^2}{b_4^2} - \varepsilon^2 b_5^2 + o(\varepsilon^2),$$

$$t_2 - t_1 = \frac{b_1\varepsilon^2 + b_2\varepsilon}{b_4} - b_5\varepsilon + o(\varepsilon) = \left(\frac{b_2}{b_4} - b_5 \right)\varepsilon + o(\varepsilon^2) + o(\varepsilon),$$

eta

$$b_1 = -\theta_2, \; b_2 = \mu_1 - \lambda_1 + \mu_2(1-\rho_1),$$

$$b_3 = \lambda_2\theta_2 + \theta_2(\mu_1-\lambda_1), \; b_4 = (\mu_2(1-\rho_1)-\lambda_2)(\mu_1-\lambda_1), \; b_5 = \frac{1}{\mu_1-\lambda_1}.$$

Hainbat kalkuluren ostean, hurrengoa lortzen da:

$$C_1(t,(\varepsilon,\varepsilon)) = \tilde{c}_1\varepsilon^2 \left(\frac{1}{2(\mu_1-\lambda_1)} \right) + \tilde{c}_2\varepsilon^2 \left(\frac{2(\mu_1-\lambda_1)+\lambda_2}{2(\mu_1-\lambda_1)^2} + \frac{(\mu_1-\lambda_1+\lambda_2)^2}{2(\mu_2(1-\rho_1)-\lambda_2)(\mu_1-\lambda_1)^2} \right) + o(\varepsilon^2).$$

Simetria dela bide, 2 klaseak lehentasuna izan balu hurrengo kostua lortuko zen

$$C_2(t,(\varepsilon,\varepsilon)) = \tilde{c}_2\varepsilon^2 \left(\frac{1}{2(\mu_2-\lambda_2)} \right) + \tilde{c}_1\varepsilon^2 \left(\frac{2(\mu_2-\lambda_2)+\lambda_1}{2(\mu_2-\lambda_2)^2} + \frac{(\mu_2-\lambda_2+\lambda_1)^2}{2(\mu_1(1-\rho_2)-\lambda_1)(\mu_2-\lambda_2)^2} \right) + o(\varepsilon^2).$$

Orain ikus daiteke $C_1(t,(\varepsilon,\varepsilon)) \leq C_2(t,(\varepsilon,\varepsilon))$ baldin eta soilik baldin $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$, (lan-karga txikiko kasua bada ($\rho < 1$)), zeinak frogapena amaitzen duen.

5.6.2 5.1. Proposizioaren frogapena

Proposizio honen frogapena PMP-tik ondoriozta daiteke, hau da, soluzio bat optimoa izan dadin bete behar diren beharrezko baldintzetatik, ikusi A.2. Eranskina, eta 5.1. Lema. Beraz, frogapenean zehar hauei egingo zaie erreferentzia.

Frogapenak hurrengo egitura jarraitzen du. PMP aplikatzen dugu muturreko soluzioak lortzeko, hau da beharrezko baldintzak betetzen dituzten soluzioak lortzeko. Lau aukera posible daudela ondorioztatuko da: $\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ indizeek orden berdina jarraitzen badute (*i.e.*, $\tilde{c}_i\mu_i \geq \tilde{c}_j\mu_j$ eta $\tilde{c}_i\mu_i/\theta_i \geq \tilde{c}_j\mu_j/\theta_j$ non $i \neq j \in \{1,2\}$) orduan lehentasuna egoera espazio osoan klase bati ematen dion politika da optimoa eta

$\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ indizeek orden ezberdina badute (*i.e.,* $\tilde{c}_i\mu_i \geq \tilde{c}_j\mu_j$ eta $\tilde{c}_j\mu_j/\theta_j \geq \tilde{c}_i\mu_i/\theta_i$ with $i \neq j \in \{1,2\}$), orduan trukatze-funtzio bat agertzen da, zeinaren azpian klase batek jasoko duen lehentasuna eta gainean beste klaseak. Orduan, $\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ indizeen ordena ezagututa bi muturreko soluzio aurkitzen dira. Ostean, 5.1. Lemak bi soluzio horien artetik onena zein den aukeratzen du.

Hipotesiz $\rho_1+\rho_2 < 1$ eta 5.3. Sekzioan aipatu den legez edozein zentzuzko politikak oreka egoera lortzen du denbora finituan, non denbora hau optimizazio parametro bat den. Orain problema honi dagozkien eta A.2. Teoreman definitu diren Hamiltondarra eta Lagrangearra idatziko dira. Hamiltondarra

$$\mathcal{H}(m(t), s(t), \gamma(t)) = \sum_{k=1}^{2} (\tilde{c}_k m_k(t) + \gamma_k(t)(\lambda_k - \mu_k s_k(t)) - \theta_k m_k(t))),$$

da eta Lagrangearra

$$\begin{aligned}\mathcal{L}(m(t), s(t), \gamma(t), \nu(t), \omega(t)) =& \mathcal{H}(m(t), s(t), \gamma(t)) - \nu_1(t)m_1(t) - \nu_2(t)m_2(t) \\ & - \omega_1(t)s_1(t) - \omega_2(t)s_2(t) + \omega_3(t)(s_1(t) + s_2(t) - 1),\end{aligned}$$

non $\gamma_k(\cdot)$ k klaseari dagokion aldagai adjuntoa den, $\nu_i(\cdot)$ Lagrange biderkatzaileak dire egoera baldintzei dagokiena, eta $\omega_i(\cdot)$ kontrol baldintzei dagokien Lagrange biderkatzaileak $i = 1, 2$ bada.

Lehenik eta behin (A.2.1). Ekuazioa ebatziko da $m_1(t), m_2(t) > 0$ den denbora tarteetan (denbora tarte hauei *barneko arkuak* deritze), zeinak $\gamma_k^*(t) = C_k' \mathrm{e}^{\theta_k t} + \frac{\tilde{c}_k}{\theta_k}$ ematen duten $k = 1, 2$ denean, eta C_k' integrazio konstanteak dira. (A.2.5). Ekuaziotik, $\nu_1(t) = \nu_2(t) = 0$ lortzen da. (A.2.3). eta (A.2.5). Ekuazioetatik $-\gamma_1(t)\mu_1 + \omega_3(t) = -\gamma_2(t)\mu_2 + \omega_3(t) = 0 \Rightarrow \mu_1\gamma_1(t) = \mu_2\gamma_2(t)$, ondorioztatzen da, $m_1(t) = 0, m_2(t) > 0$ edo $m_2(t) = 0, m_1(t) > 0$ betetzen den edozein denbora tartetarako, tarte hauek *mugako arku* izenez ezagutzen dira. Mugako arkuetan $m_k(t) = 0$ denez, $\mathrm{d}m_k(t)/\mathrm{d}t = 0$ eta beraz $s_k(t) = \lambda_k/\mu_k$ eta $\omega_k(t) = 0$ $k = 1, 2$ bada.

Orain (A.2.2). Ekuazioa ebatziko da barneko arkuetan, zeina

$$\arg\min_{\mathbf{s}\in\mathcal{S}} \sum_{k=1}^{2} -\mu_k\gamma_k(t)s_k(t) = \arg\min_{\mathbf{s}\in\mathcal{S}} \sum_{k=1}^{2} -\mu_k \left(C_k' \mathrm{e}^{\theta_k t} + \frac{\tilde{c}_k}{\theta_k}\right) s_k(t), \tag{5.6.2}$$

ekuazioa ebaztearen baliokidea den. Azken honek, kontrol optimoak $\mu_k \left(C_k' \mathrm{e}^{\theta_k t} + \frac{\tilde{c}_k}{\theta_k}\right)$ baliorik altuena duen klaseari emango diola lehentasuna inplikatzen du branch arkuetan. C_k integrazio konstanteak ezagunak direnean $k = 1, 2$-rentzat Lagrange biderkatzaileak eta bektore adjuntoak guztiz karakteriza daitezke.

Beraz, C_k' konstanteak kalkulatuko dira $k = 1, 2$ denean, zeinak lehentasunak ezartzen lagunduko duen. Lehenik eta behin, lehentasunaren aldekatarik gertatzen den aztertuko da hurrengo trukatze-funtzioa aztertuz:

$$\sigma(t) := \mu_1 \left(C_1' \mathrm{e}^{\theta_1 t} + \frac{\tilde{c}_1}{\theta_1}\right) - \mu_2 \left(C_2' \mathrm{e}^{\theta_2 t} + \frac{\tilde{c}_2}{\theta_2}\right).$$

Ohartu $\sigma(t)$ funtzioak gehienez bi erro dituela, muturreko soluzioa beraz hurrengo soluzioetatik bat izango da:

(i) Soluzioa lehentasun hertsia da, hau da, $\nexists\, t \in [0, T]$ s.t $\sigma(t) = 0$.

(ii) Trukatze-funtzio bat existitzen da zeinaren gainetik klase batek duen lehentasuna eta azpitik besteak, hau da, $\exists!\, t_1 \in [0, T],$ non $\sigma(t_1) = 0$.

(iii) Soluzioa jatorritik gertu klase bati lehentasuna ematea da, jatorritik urruti klase berdin horri eta erdian beste klaseari, hau da, $\exists\, t_1, t_2 \in [0, T]$ zeinarentzat $t_1 > t_2$ eta $\sigma(t_1) = \sigma(t_2) = 0$.

Orokortasunik galdu gabe $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ dela onartu da. Beste kasua argumentazio bera erabiliz azter daiteke. PMP aplikatuko dugu (iii) soluzioa ez dela aukera posiblea ikusteko, eta (i) eta (ii) posible izan daitezen baldintzak lortuko dira.

Hori egin ahal izateko, lehenik eta behin trukaze-funtzio bat existitzen dela onartuko da. Notazioa erreztearren hasierako puntua $(m_1(0), m_2(0)) = (m_{10}, m_{20})$, $m_{10}, m_{20} > 0$ trukatze-funtzioan dagoen puntu bat dela onartuko da. Bi trukatze-funtzio dauden kasuan (m_{10}, m_{20}), $(0, 0)$ oreka egoera iritsi aurreko trukatze-funtzioan dagoela onartuko da. Gogoratu orain $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$, hipotesia, beraz 5.1. Lema dela-eta $s^*(t) = (0, 1)$ da $t_1 = 0$ bada, eta $s^*(t) = (1, 0)$ $t \in (t_1, t_2]$ bada, non t_2 ibilbidea mugako arkuan sartzen den unea den $m_1^*(t) = 0$ eta $s^*(\rho_1, 1-\rho_1)$ edozein $t \in (t_2, t_3]$ bada, non t_3 oreka egoera ikutzen den unea den.

Orain trukatze-funtzioa aztertuko da, zeinak zein klasek jasotzen duen lehentasuna karakterizatzen duen. C_1' eta C_2' konstanteak lortzeko, PMP-ko (A.2.6). Ekuazioa erabiliko da.

$t = t_1 = 0$ bada, orduan $s_2^*(t) = 1$ eta $s_1^*(t) = 0$. $\gamma_k^*(t) = C_k' \mathrm{e}^{\theta_k t} + \tilde{c}_k/\theta_k$, dela erabiliz, Hamiltondarra $t = t_1 = 0$ unean

$$\mathcal{H}(m^*(t), s^*(t), \gamma^*(t), t) = \tilde{c}_1 \frac{\lambda_1}{\theta_1} - C_1'\theta_1 \left(m_{10} - \frac{\lambda_1}{\theta_1} \right) + \tilde{c}_2 \left(\frac{\lambda_2 - \mu_2}{\theta_2} \right) - \theta_2 C_2' \left(m_{20} + \frac{\mu_2 - \lambda_2}{\theta_2} \right), \tag{5.6.3}$$

da. $t \in (0, t_2]$ bada, orduan $s_1^*(t) = 1$ eta $s_2^*(t) = 0$. Beraz,

$$m_1^*(t) = \left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1} \right) \mathrm{e}^{-\theta_1 t} + \frac{\lambda_1 - \mu_1}{\theta_1}, \qquad m_2^*(t) = \left(m_{20} - \frac{\lambda_2}{\theta_2} \right) \mathrm{e}^{-\theta_2 t} + \frac{\lambda_2}{\theta_2},$$

horrela $t \in (0, t_2]$ bada

$$\begin{aligned}
&\mathcal{H}(m^*(t), s^*(t), \gamma^*(t), t) \\
&= \tilde{c}_1 \left(\left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1} \right) \mathrm{e}^{-\theta_1 t} + \frac{\lambda_1 - \mu_1}{\theta_1} \right) + \tilde{c}_2 \left(\left(m_{20} - \frac{\lambda_2}{\theta_2} \right) \mathrm{e}^{-\theta_2 t} + \frac{\lambda_2}{\theta_2} \right) \\
&\quad + \left(C_1' \mathrm{e}^{\theta_1 t} + \frac{\tilde{c}_1}{\theta_1} \right) \left(\lambda_1 - \mu_1 - \theta_1 \left(\left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1} \right) \mathrm{e}^{-\theta_1 t} + \frac{\lambda_1 - \mu_1}{\theta_1} \right) \right) \\
&\quad + \left(C_2' \mathrm{e}^{\theta_2 t} + \frac{\tilde{c}_2}{\theta_2} \right) \left(\lambda_2 - \theta_2 \left(\left(m_{20} - \frac{\lambda_2}{\theta_2} \right) \mathrm{e}^{-\theta_2 t} + \frac{\lambda_2}{\theta_2} \right) \right) \\
&= \tilde{c}_1 \left(\frac{\lambda_1 - \mu_1}{\theta_1} \right) + \tilde{c}_2 \frac{\lambda_2}{\theta_2} - \theta_1 C_1' \left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1} \right) - \theta_2 C_2' \left(m_{20} - \frac{\lambda_2}{\theta_2} \right).
\end{aligned} \tag{5.6.4}$$

(5.6.3). eta (5.6.4). Ekuazioak 0-z berdinduz, hurrengo espresioa lortzen da:

$$C_1' = \frac{\tilde{c}_1 \left(\frac{\lambda_1 - \mu_1}{\theta_1} \right) + \tilde{c}_2 \frac{\lambda_2}{\theta_2} - \theta_2 C_2' \left(m_{20} - \frac{\lambda_2}{\theta_2} \right)}{\theta_1 \left(m_{10} + \frac{\mu_1 - \lambda_1}{\theta_1} \right)},$$

$$C_2' = \frac{\left(\frac{\tilde{c}_1\mu_1}{\theta_1} - \frac{\tilde{c}_2\mu_2}{\theta_2} \right) \theta_1 m_{10} - \mu_1 \frac{\tilde{c}_2\mu_2}{\theta_2} (1 - \rho_1 - \rho_2)}{\mu_1\theta_2 m_{20} + \theta_1\mu_2 m_{10} + \mu_1\mu_2(1 - \rho_1 - \rho_2)}. \tag{5.6.5}$$

$t \in (t_2, t_3]$ bada, orduan $s_1^*(t) = \rho_1$, goian kalkulatu den legez, eta $s_2^*(t) = 1 - \rho_1$. Beraz,

$$m_1^*(t) = 0,$$

$$m_2^*(t) = \left(m_{20} - \frac{\lambda_2}{\theta_2} + \frac{\mu_2\mu_1 - \lambda_1\mu_2}{\mu_1\theta_2} \left(\frac{\mu_1 - \lambda_1}{m_{10}\theta_1 - \lambda_1 + \mu_1} \right)^{-\dfrac{\theta_2}{\theta_1}} \right) \mathrm{e}^{-\theta_2 t} - \left(\frac{\mu_2\mu_1 - \lambda_1\mu_2 - \lambda_2\mu_1}{\mu_1\theta_2} \right),$$

eta $t \in (t_2, t_3]$ bada orduan

$$\begin{aligned} \mathcal{H}(m^*(t), s^*(t), \gamma^*(t), t) &= \tilde{c}_2 m_2^*(t) + \left(C_2' \mathrm{e}^{\theta_2 t} + \frac{\tilde{c}_2}{\theta_2} \right) \left(\lambda_2 - \mu_2 \left(1 - \frac{\lambda_1}{\mu_1} \right) - \theta_2 m_2^*(t) \right) \\ &= -\frac{\tilde{c}_2\mu_2}{\theta_2}(1 - \rho_1 - \rho_2) \\ &\quad -\theta_2 C_2' \left(m_{20} - \frac{\lambda_2}{\theta_2} + \frac{\mu_2\mu_1 - \lambda_1\mu_2}{\mu_1\theta_2} \left(\frac{\mu_1 - \lambda_1}{m_{10} - \lambda_1 + \mu_1} \right)^{-\dfrac{\theta_2}{\theta_1}} \right). \end{aligned} \tag{5.6.6}$$

(5.6.6). Ekauzioa 0-rekin berdinduz, eta (5.6.5). Ekuazioko C_1' eta C_2' konstanteen espresioak ordezkatuz, jatorritik gertuen dagoen trukatze-funtzioak hurrengo erlazioa betetzen du:

$$m_{20} = \frac{a_1 m_{10} + a_2 + (a_3 m_{10} - a_2) \left(\dfrac{\theta_1 m_{10} + \mu_1 - \lambda_1}{\mu_1 - \lambda_1} \right)^{\dfrac{\theta_2}{\theta_1}}}{a_4 m_{10}} + \frac{\lambda_2}{\theta_2}, \tag{5.6.7}$$

non

$$a_1 = \tilde{c}_2 \frac{\mu_2}{\theta_2} (1 - \rho_1 - \rho_2); \quad a_2 = a_1 \frac{\mu_1}{\theta_1}(1 - \rho_1);$$

$$a_3 = \left(\tilde{c}_1 \frac{\mu_1}{\theta_1} - \tilde{c}_2 \frac{\mu_2}{\theta_2} \right) (1 - \rho_1); \quad a_4 = -\left(\tilde{c}_1 \frac{\mu_1}{\theta_1} - \tilde{c}_2 \frac{\mu_2}{\theta_2} \right) \frac{\theta_2}{\mu_2},$$

diren. Azken espresio honek trukatze-funtzio bat definitzen du lehenengo kuadrantean $m_{20} > 0$ bada $m_{10} \to 0$ den heinean. L'Hopital erabiliz

$$m_{20} \xrightarrow{m_{10} \to 0} \frac{a_1}{a_4} + \frac{\lambda_2}{\theta_2} + \frac{a_3}{a_4} - \frac{a_2\theta_2}{a_4\mu_1(1 - \rho_1)} = (1 - \rho_1 - \rho_2) \frac{\mu_2}{\theta_1\theta_2} \left(\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\dfrac{\tilde{c}_2\mu_2}{\theta_2} - \dfrac{\tilde{c}_1\mu_1}{\theta_1}} \right),$$

lor daiteke. Sistemak lan-karga txikia duenez ($\rho_1 + \rho_2 < 1$), eta orokortasunik galdu gabe $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ onartu denez

$$(1 - \rho_1 - \rho_2) \frac{\mu_2}{\theta_1\theta_2} \left(\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\dfrac{\tilde{c}_2\mu_2}{\theta_2} - \dfrac{\tilde{c}_1\mu_1}{\theta_1}} \right) \geq 0 \Longleftrightarrow \tilde{c}_2\mu_2/\theta_2 > \tilde{c}_1\mu_1/\theta_1 \text{ and } \tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2,$$

da. Berdina lor daiteke indizeak aldatuta $\tilde{c}_2\mu_2 \geq \tilde{c}_1\mu_1$ onartuko balitz. Orduan, lehentasunen aldaketa bat gerta dadin (aldaketa bat gutxienez) $\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ indizeek aderantzizko ordena izan behar dute. Beraz, frogatu da $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu$ orden berdinekoak badira, ez dagoela trukatze-funtziorik. Azken honek 5.1.

Lemarekin batera $\tilde{c}\mu$ indizerik handiena duen klaseak duela lehentasuna frogatu da (edo baliokideki $\tilde{c}\mu/\theta$ handiena duena). $\tilde{c}\mu$ eta $\tilde{c}\mu/\theta$ indizeek orden berdina duten kasurako politika lortu da, orain $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu$ indizeek orden ezberdina duten kasua aztertu behar da. Azken honetan (ii) eta (iii) dira aukera posibleak, ikusiko da (iii) ez dela inoiz gertatzen.

Izan bedi $\tilde{c}_2\mu_2/\theta_2 > \tilde{c}_1\mu_1/\theta_1$ eta $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$. Bi trukatze-funtzio existitzen direla onartuko da. Eta onartu hasierako puntua jatorritik gertuen dagoen funtzioaren gainean daogela, hau da, goian karakterizatu den trukatze-funtzioan. Orduan, $\sigma(0) = 0$ da eta C_1' eta C_2' (5.6.5). Ekuazioak definitzen ditu. (m_{10}, m_{20}) hasiera puntua trukatze funtzioaren (SC-ren) gainean dagoela onartu izan ez balitz, eta ordez, jatorritik urrutien dagoen trukatze-funtzioaren gainean dagoela onartu balitz, orduan existituko litzateke $t' \in (0, T]$ zeinarentzat $\sigma(0) = \sigma(t') = 0$. Amaierako denbora optimizazioak erabakitzen duenez $[-t', T]$ tartea kontsideratu daiteke, eta antzeko moduan argudiatuz, hasierako puntua lehenengo trukatze funtzioan aurkitzeak $-t$ denbora bigarren trukatze funtzioa $t = 0$ unean aurkitzea inplikatzen du. Orduan, bigarren trukatze-funtzio bat existitu dadin $t' > 0$ existitzen da zeinarentzat $\sigma(-t') = 0$ den. Halere, behean ikusiko da $\sigma(t)$ hertsiki gorakorra dela $t < 0$ denean, eta beraz bigarren trukatzea ez da inoiz gertatzen.

$\sigma(t)$ hertsiki gorakorra dela frogatzeko edozein $t < 0$ bada, $\mathrm{d}\sigma(t)/\mathrm{d}t > 0$ dela ikusiko da $t < 0$ denean. (5.6.5). Ekuaziotik $C_2' \leq 0$ da edozein $m_{20}, m_{10} > 0$ denean, orduan

$$\theta_1\mu_1C_1'\mathrm{e}^{\theta_1 t} - \theta_2\mu_2C_2'\mathrm{e}^{\theta_2 t} > 0 \Leftrightarrow \frac{\theta_1\mu_1C_1'}{\theta_2\mu_2C_2'} < \mathrm{e}^{(\theta_2-\theta_1)t}. \tag{5.6.8}$$

C_1' eta C_2' konstanteen espresioak ordezkatuz eta kalkulu batzuen ostean

$$\frac{\theta_1\mu_1C_1'}{\theta_2\mu_2C_2'} \leq 1 \Leftrightarrow m_{20} \geq (1-\rho_1-\rho_2)\frac{\mu_2}{\theta_1\theta_2}\left(\frac{\tilde{c}_1\mu_1 - \tilde{c}_2\mu_2}{\dfrac{\tilde{c}_2\mu_2}{\theta_2} - \dfrac{\tilde{c}_1\mu_1}{\theta_1}}\right) - \frac{\mu_2}{\mu_1}m_{10},$$

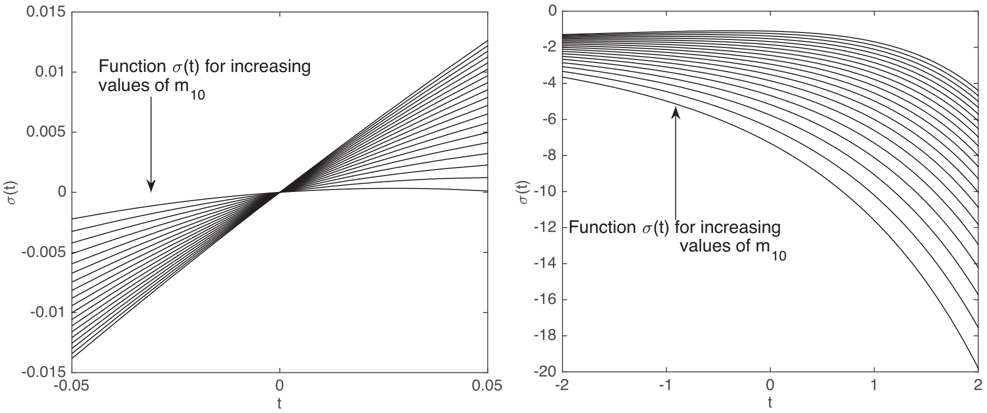
lortzen da. Nahikoa da orain (5.6.7). Ekuazioko m_{20}-ren espresioa ordezkatzea eta azken inekuazioa frogatzea. Ohartu inekuazioaren eskuinaldea $\dfrac{a_1}{a_4} + \dfrac{\lambda_2}{\theta_2} + \dfrac{a_3}{a_4} - \dfrac{a_2\theta_2}{a_4\mu_1(1-\rho_1)} - \mu_2 m_{10}/\mu_1$ dela eta $-\mu_2/\mu_1 = a_3\theta_2/(a_4(\mu_1-\lambda_1))$. Orduan hainbat kalkuluren ostean

$$\begin{aligned}&\frac{\frac{\mu_1}{\theta_1}\frac{\tilde{c}_2\mu_2}{\theta_2}(1-\rho_1-\rho_2)}{\frac{\tilde{c}_1\mu_1}{\theta_1} - \frac{\tilde{c}_2\mu_2}{\theta_2}}\left(1 + \frac{\theta_2 m_{10}}{\mu_1-\lambda_1} - \left(\frac{\theta_1 m_{10} + \mu_1 - \lambda_1}{\mu_1-\lambda_1}\right)^{\frac{\theta_2}{\theta_1}}\right)\\ &\leq m_{10}\left(1 + \frac{\theta_2 m_{10}}{\mu_1-\lambda_1} - \left(\frac{\theta_1 m_{10} + \mu_1 - \lambda_1}{\mu_1-\lambda_1}\right)^{\frac{\theta_2}{\theta_1}}\right),\end{aligned} \tag{5.6.9}$$

lortzen da, izan ere, $\tilde{c}_2\mu_2/\theta_2 > \tilde{c}_1\mu_1/\theta_1$ eta $m_{10} \geq 0$. Definitu

$$\ell(m_{10}) := \left(1 + \frac{\theta_2 m_{10}}{\mu_1-\lambda_1} - \left(\frac{\theta_1 m_{10} + \mu_1 - \lambda_1}{\mu_1-\lambda_1}\right)^{\frac{\theta_2}{\theta_1}}\right).$$

Orduan, (5.6.9). Inekuazioa bete dadin nahiko a da $\ell(m_{10}) \geq 0$ dela frogatzea edozein $m_{10} \geq 0$ bada. Ohartu $\ell(0) = 0$ eta $\ell(\cdot)$ ez-beherakorra dela edozein $m_{10} > 0$, beraz, $\ell(m_{10}) \geq 0$ edozein $m_{10} \geq 0$ bada.



Irudia 5.11: Ezkerraldea: $\sigma(t)$ trukatze funtzioa $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu$ indizeek orden ezberdina jarraitzen dutenean. Eskuinaldea: $\sigma(t)$ trukatze-funtzioa $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu$ indizeek orden berdina dutenean.

Azken honek frogatzen du $\theta_1\mu_1 C_1'/(\theta_2\mu_2 C_2') \leq 1$ eta beraz $\sigma(t)$ hertsiki gorakorra da edozein $t < 0$ denean, zeinak inplikatzen duen ez dela trukatzerik gertatzen.

Orain, $\tilde{c}_1\mu_1 \geq \tilde{c}_2\mu_2$ dela onartu denez, eta $\tilde{c}_2\mu_2/\theta_2 > \tilde{c}_1\mu_1/\theta_1$ $\theta_1 > \theta_2$ da. Azken honek zera inplikatzen du $e^{(\theta_2-\theta_1)t} > 1$ edozein $t < 0$, beraz, (5.6.8). Ekuazioa edozein $t < 0$-rako betetzen da, ikusi 5.11. Irudia, non $\sigma(t)$ funtzioak erro bakarra duela ikus daitekeen $\tilde{c}\mu/\theta$ eta $\tilde{c}\mu$ indizeek aurkako ordena dutenean.

Atala III

Sistema estokastikoen kontrol dinamiko optimoa

6. Kapitulua

Uzteak gerta daitezkeen sorta-ilarak

Contents

6.1 Sarrera **132**

6.2 Ereduaren deskribapena **133**

6.3 Atari-politika optimoak **135**

6.4 Banaketa egoera-egonkorrean **135**

6.5 Atari-politika optimoaren karakterizazioa **138**

6.6 Adibideak **139**

6.7 Eranskina **142**

Kapitulu honetan edukien banaketarako problema aztertu da bezeroak sortatan zerbitzatzen direnean eta beraien zerbitzua hasi aurretik sistema utz dekatekenean. Problema zerbitzari bakarreko ilara Markovtarra bezela irudikatuko da eta hurrengo bi kasuak aztertuko dira: (1) zerbitzua aktibatu bezain laster bezeroak zerbitzatuak izango dira, *i.e.*, zerbitzu tasa infinitua da, eta (2) zerbitzu tasa finitua den kasua. Helburua batez besteko mantentze-kostua minimizatzen duen politika aurkitzea da, bezeroak zerbitzua uztean penalizazio bat ordaintzen delarik eta zerbitzua aktibatzen den bakoitzean aparteko kostu bat ordaintzen delarik. Azken kostu hau zerbitzua aktibatzen denero ordaindu beharrekoa da zerbitzu tasa infinitua denean, eta zerbitzaria aktiboa den denbora unitatero zerbitzua finitua den kasuan. Uzteen perspektibatik begiratuz desigaragarria da sorta txikietan zerbitzatzea bezeroak; baina zerbitzuak kostua handia suposatzen duenez hobe da sorta handiagotan elkartzea bezeroak. Kapitulu honen helburua bi kontzeptu horien arteko oreka aurkitzea da.

Aurreko kapituluetan soluzio optimoaren hurbilketa onak diren soluzioak aurkitzen saiatu gara, bertako ereduen zailtasuna dela-eta. Kapitulu honetan ordea ilara klase bakarra denek soluzio optimoa esplizituki karakterizatzea ahalbideratzen du.

6.1 Sarrera

Kapitulu honetan sorta ilarak eta bezeroen uzteak bateratzen dituen sistema bat aztertuko da. Sistema hau M/M/1 ilara bat da zeinak zerbitzu pozesu berezi bat duen, bezeroen zerbitzua atzeratu daiteke beste hainbat bezeroekin batera (sortan) zerbitzatua izan dadin. Bi egoera ezberdin kontsideratuko dira: (1) bezeroak, sorta bat zerbitzuan hartu bezain laster prozesatuak dira, eta sistematik badoaz, (2) zerbitzu denborak banaketa esponentziala du, zeinak $1/\mu$ batez-bestekoa duen. Sorta baten zerbitzu denbora sorta horretako bezero kopuruarekiko askea da (*multi-cast* izenez ezagutzen da). Sortak sortzeko bezeroen zerbitzua atzeratzean bezeroen uzteak eragin ditzake. Bereziki, bezeroek sistema utz dezakete zerbitzua jasotzeko zain daudenean, zeinagatik sistemak penalizazio bat jasotzen duen kostu finko batekoa sistema utzi duen bezero bakoitzeko. Penalizazio hauek bezeroak galtzeagatik ordain daitezke edo bezero horiek *back-up* zerbitzu batean zerbitzatzeagatik. Uzteen prozesua, bezero bakoitzaren denbora amaiera, banaketa esponentzial bidez irudikatu da. Antzeko metodo bat kontsideratu da [58] artikuluan, zeinen autoreek uzteak gertatzen diren sistema bat aztertzen duten sorta zerbitzuak kontsideratuz, baina beraien soluzioak bezeroak banaka zerbitzatu behar direla ondorioztatzen dute.

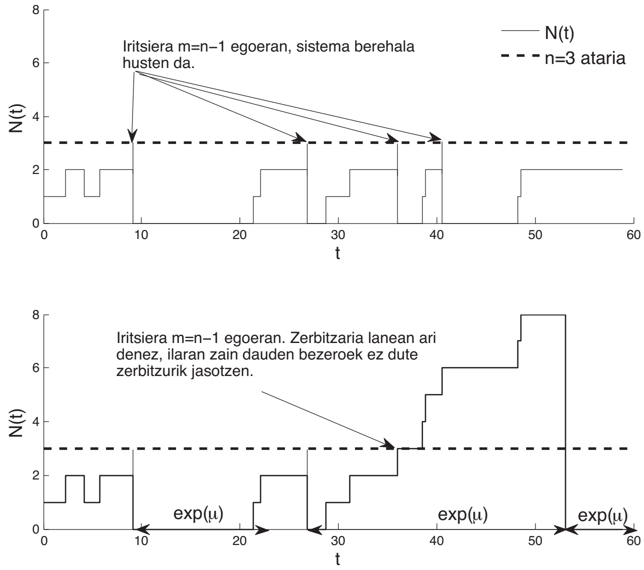
Multi-cast trafikoaren antolamendua bezeroen uzteekin hainbat aplikaziotan aurki daiteke, *e.g.*, haririk gabeko sentsore sareetan, telefono mugikorretan bideoak ikuskatzea. Ikerlari askok aztertu dituzke epemugak dituzten sistemak, zeinetarako bezeroen zerbitzu bat jasotzeko epemugak ezagunak diren. Hauek oso aztertuak izan diren antolamendu problemak dira, horrelako ilara eta sareentzat *Earliest Deadline First (EDF)* politika ezaguna da. Adibidez, EDP ilarak lar karga handietarako [41] eta [64] artikuluetan aztertua izan da. EDF optimoa izatea [88] artikuluan aztertu da, non epemuga amaitu zaien bezeroen kopurua minimizatzea zen helburua, eta zerbitzuak banaketa esponentziala jarraitzen zuen. EDF politikak eta antzeko politikek bezero bakoitzarentzat zerbitzu bat dagoela onartzen dute, ordez tesi honetan bezeroak elkartu daitezkeela onartu da.

Lehenengo ekarpena politikak monotonoak direla frogatzea da, zeinak kasu honetan atari-politikak diren. Egitura emaitzak adibidez [74] artikuluan lortu dira sorta zerbitzuak dituen sistema baterako baina edukiera finituko zerbitzari batetarako eta uzterik gabeko sistema baterako.

Ilaretan uzteak izateak sistema ez-uniformizagarria izatea eragiten du, eta beraz, politika monotonoak direla soluzio optimoa frogatzea zaila gertatzen da. Arazo honi aurre egiteko SRT metodoa proposatu da [25] artikuluan, irakurlea 1.3.3. Sekziora joan daiteke uniformizaziori buruzko teknikak ikasteko.

Bigarren ekarpenean sistemaren jokabidea aztertzen da egoera egonkorrean eta egoera-egonkorreko ilara-luzeera kalkulatuko da edozein $\mu \in \mathbb{R}_+ \cup \{\infty\}$ bada, horretarako funtzio sortzaileen hurbilketa erabiliko da. Azken honek zerbitzua hasteagatik ordaindu beharreko kostua aurkitzea ahalbideratzen du, sistemaren edozein egoera posibletan, zeinetarako sorta bat zerbitzuan hartzea eta zerbitzaria lanik egin gabe uztea sistemarentzat berdinak diren kostuari dagokionean. Zerbitzariaren abiadura infinitua denean, ataria guztiz karakteriza daiteke. Zerbitzu abiadura finitua denean, analisia zailagoa da eta karakterizazio oso bat ezin izan da aurkitu, halere, soluzio optimoa optimizazio baten emaitza den atari politika batek ematen du.

Kapitulo honen gainontzekoak hurrengo egitura jarraitzen du. 6.2. Sekzioan eredua deskribatu da. 6.3. Sekzioan atari-politikak optimoak direla frogatu da. 6.4. Sekzioan egoera-egonkorrean ilara luzeeraren banaketa aurkitu da, zeinak 6.5. Sekzioan atari-politika optimoa karakterizatzea ahalbideratzen duen. Azkenik, 6.6. Sekzioan adibide numerikoen bitartez lortu diren emaitzak irudikatuko dira. Frogapen gehienak 6.7. Eranskinean aurki daitezke.



Irudia 6.1: $N(t)$ prozesuaren simulazioa, non $N(t)$ t denboran ilaran zain dauden bezero kopurua den, $n = 3$ atari-politikapean. Goian $\mu = \infty$ kasua, behean $\mu < \infty$ kasua. Behean $exp(\mu)$ zerbitzariaren denbora lanpetuaren adierazpide den, denbora honek banaketa esponentziala jarraitzen du $\mu < \infty$ parametroarekin. Honen ondorioz $N(t)$ n atariaren menpekoa da baina baita lanpetutako denboraldien luzeeraren araberakoa.

6.2 Ereduaren deskribapena

M/M/1 ilara bat kontsideratu da sorta zerbitzuekin, edukiera infinitua eta bezeroak sistema utz dezakete. Bezeroak ilarara λ parametroko Poisson prozesu bat jarraituz iristen dira eta banaketa esponentziala duen zerbitzua jasotzen dute, $1/\mu$ batez bestekoarekin, zeina sortaren tamainarekiko askea den. Ilaran zain dauden bezeroak sistema utz dezaketa θ parametroko banaketa esponentziala jarraitzen duen denbora tartearen ostean. Ez hori bakarrik, iritsiera batetik besterako tarteak, zerbitzu denborak eta uzteak askeak dira.

Uneoro ϕ politikak erabakitzen du zain dauden bezeroak zerbitzatu ala ez. Bezero bat zerbitzura onartu den bezain laster ezingo du sistema utzi. Izan bedi $N^\phi(t) \in \{0, 1, \ldots\}$ ilaran zain dauden bezeroen kopurua t denboran eta ϕ politikapean. Izan bedi $S^\phi(N^\phi(t)) \in \{0, 1\}$, t denboran eta ϕ politikapean sistemak hartutako erabakiaren adierazgarri, non $N^\phi(t)$ sisteman zain dauden bezero kopurua den. Hau da, $S^\phi(N^\phi(t)) = 0$ zerbitzaria lanean ari ez bada, eta $S^\phi(N^\phi(t)) = 1$ zerbitzaria sorta bat zerbitzatzen ari bada. Zerbitzariaren edukiera infinitua dela-eta, zerbitzua aktibatu bezain laster, *i.e.*, $S^\phi(N^\phi)(t) = 1$, ilaran zain dauden bezero guztiak zerbitzua hasiko dute. Beraz, sorta baten luzera, zerbitzua aktibatzen den uneko zain dauden bezeroen tamainakoa izango da.

Problema hau bi egoera ezberdinetan aztertuko da (ikusi 6.1. Irudia, non ϕ atari-politika dela kontsideratu den):

- Sistema sorta bat zerbitzuan hartzeko erabakia hartu bezain laster husten da, hau da, $\mu = \infty$.
- Zerbitzuak banaketa esponentziala jarraitzen du $\mu < \infty$ tasarekin.

($\mu = \infty$) kasuan, zerbitzua berehalakoa da, zerbitzariak sistema husten du zerbitzaria aktibatzeko erabakia hartu bezain laster. Honek aiderazten du ϕ politikak zerbitzua aktibatzea erabakitzen duenero, $N^{\phi}(t)$ luzerako sorta bat (*i.e.,* ilaran zain dauden bezero guztiak) berehala prozesatuak izango dira. Bigarren kasuan ($\mu < \infty$), zerbitzaria aktibatzen denean zerbitzariak $N^{\phi}(t)$ tamainuko sorta bat hartze du zerbitzura, eta banaketa esponentziala duen denboraldi bat eskaintzen du sorta hori prozesatzeko. Zerbitzaria lanpetuta dagoen bitartean, bezero berriak irits daitezke ilarara. Kasu honetan, zerbitzariak ezin du sorta berri bat hartu zerbitzuan aurreko sortaren zerbitzua amaitu den arte; ikusi 6.1. Irudia (behean) $t = 37$ inguruan. Beraz, $N^{\phi}(t)$-ren eboluzioa ϕ politikaren eta zerbitzuaren egoeraren araberakoa da (*i.e.*, lanpetua ala aske). $\mu = \infty$ kasuan, sistemaren egoera ilaran zain dauden bezero kopurua da. $\mu < \infty$ kasuan ordea, sistemaren egoera (m, a) da, non m ilaran zain dauden bezeroen kopurua den eta $a \in \{0, 1\}$-k zerbitzaria lanpetua ($a = 1$) edo aske dagoen ($a = 0$) adierazten duen.

c-ren bidez bezero bakoitza ilaran mantentzeagatik ordaindu beharreko denbora unitateko kostua adierazten da, δ bezeroek ilara uztean eragiten den kostua da, c_s^{∞} zerbitzaria aktibatzen den bakoitzeko ordaindu beharreko kostua da $\mu = \infty$ kasuan eta c_s zerbitzaria lanean ari den denbora unitateko ordaindu beharreko kostua $\mu < \infty$ kasuan. Lan honen helburua ϕ politika aurkitzea da zeinak hurrengo kostua minimizatzen duen

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\bigg[\int_0^T \left[cN^{\phi}(t) + c_s^{\infty}\lambda \mathbf{1}_{\{N^{\phi}(t)\in\mathcal{M}^{\phi}\}}\right] \mathrm{d}t + \delta R^{\phi}(T)\bigg],$$

$\mu = \infty$ kasuan eta

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\bigg[\int_0^T \left[cN^{\phi}(t) + c_s S^{\phi}(N^{\phi}(t))\right] \mathrm{d}t + \delta R^{\phi}(T)\bigg],$$

$\mu < \infty$ kasuan. Azken funtzio objetiboan $R^{\phi}(T)$ $[0, T]$ denbora tartean sistema utzi duten bezeroen kopuruaren adierazle da, eta $\mathcal{M}^{\phi} = \{m \in \{0, 1, \ldots\} : S^{\phi}(m) = 0, S^{\phi}(m+1) = 1\}$. Dynkin-en formula erabiliz [3] hurrengoa ondorioztatzen da: $\mathbb{E}[R^{\phi}(T)] = \theta\mathbb{E}[\int_0^T N^{\phi}(t)\mathrm{d}t]$, eta beraz azken funtzio objetiboa hurrengo funtzioaren baliokidea da: aurkitu ϕ zeinak

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T [\tilde{c}N^{\phi}(t) + c_s^{\infty}\lambda \mathbf{1}_{\{N^{\phi}(t)\in\mathcal{M}^{\phi}\}}]\mathrm{d}t\right],$$

minimizatzen duen $\mu = \infty$ kasuan, eta

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T [\tilde{c}N^{\phi}(t) + c_s S^{\phi}(N^{\phi}(t))]\mathrm{d}t\right],$$

$\mu < \infty$ kasuan, non $\tilde{c} = c + \delta\theta$. Sistema ergodikoa denez, denboran batezbeste optimoa den politika egoera-egonkorrean optimoa den politikaren baliokidea da, eta beraz helburua ϕ aurkitzea da zeinetarako

$$\min_{\phi} \left(\tilde{c}\mathbb{E}[N^{\phi}] + c_s^{\infty}\lambda\mathbb{E}(\mathbf{1}_{\{N^{\phi}\in\mathcal{M}^{\phi}\}})\right),$$

den $\mu = \infty$ kasuan, eta

$$\min_{\phi} \left(\tilde{c}\mathbb{E}[N^{\phi}] + c_s\mathbb{E}(\mathbf{1}_{\{S^{\phi}(N^{\phi})=1\}})\right),$$

$\mu < \infty$ kasuan. Goian deskribatu den problema MEP bat da eta P problema bezela deituko zaio kapitulo honetan zehar.

Orokorrean problema hauek oso zailak gerta daitezke ebazterakoan problemen dimentsioa dela-eta eta transizioak ez bornatuak direnez sistema ez da uniformizagarria. 6.3. Sekzioan ikusiko den legez [25] artiluloko metodoak P problemaren egiturari dagokion emaitzak lortzea ahalbideratzen du. Ikusiko da bi kasuetan, $\mu = \infty$ eta $\mu < \infty$, atari-politikak direla batezbeste optimoak. Hau da, existitzen dela n atari bat zeinentzat zerbitzariak ez dituen bezeroak zerbitzatuko edozein $m \leq n-1$ bada eta zerbitzuan hartuko ditu bezeroak bestela.

6.3 Atari-politika optimoak

Egitura propietateak lortzea uniformizagarriak ez diren problementzat nahiko zaila gerta daiteke. 3. Kapituluan ikusi da zenbait baldintza betetzen badira atari-politikak optimoak direla ikusi daitekeela SRT metodoa erabiliz, metodo hau [25] artikuluan aurkeztu da, ikusi baita 1.3.3. Sekzioa eztabaida labur batentzat. Sekzio honetan atari-politikak optimoak direla ikusiko da metodo hauxe erabiliz. Horretarako SRT aplikatu ahal izateko baldintzak betetzen direla ikusiko da lehenik eta behin.

Lehenik eta behin egoera espazio finituko MEP-a definituko da, hau da, izan bedi $L < \infty$ eta ilaran zain dauden bezero kopurua $m \in \{0, \ldots, L\}$. Ez hori bakarrik, iritsiera tasa hurrengo moduan moldatuko da: $q^{\phi,L}(m, m+1) = \lambda\left(1 - \frac{m}{L}\right)$, edozein $L \geq m \geq 0$ denean. Trunkaketa honek transizio tasa bornatuak lortzea ahalbideratzen du, eta iritsiera tasen moldaketak jatorrizko problemaren egitura mantentzea ahalbideratzen du.

MEP-ren egoera espazio finitua definituko da, eta 6.7.1. Eranskinean frogatuko da [25, 3.1. Teorema] teoremako baldintzak betetzen dituela P problemak, zeinak SRT metodoa erabilgarria izatea ahalbideratzen duen.

Ondoren atari-politikak optimoak direla frogatuko da $\mu \in \mathbb{R}^+ \cup \{\infty\}$. Frogapena $\mu = \infty$ kasuan 2.1. Proposizioan aurkitzen da, eta $\mu < \infty$ kasuan 6.7.1. Eranskinean aurki daiteke [25, 3.1. Teorema] teoremako baldintzen ziurtatzearekin batera.

6.1 Proposizioa. *$\mu \in \mathbb{R}^+ \cup \{\infty\}$ izanik, existitzen da n atari-politika bat P problemarentzat optimoa.*

6.4 Banaketa egoera-egonkorrean

Sekzio honetan P problemaren egoera egonkorreko banaketa kalkulatuko da atari politikapean. Politika hauek n atari edo $\phi = n$ gisa izendatuko dira, ataria $n \in \{0, 1, \ldots\}$ denean. Gogoratu kapitulu honetan ataria n izatean akzio pasiboa inplikatzen duela $m \leq n-1$, egoeretan eta aktiboa $m \geq n$ egoeretan. Azalpena erraztearren notaziotik n-ren dependentzia kenduko da.

6.4.1. Sekzioan P problemaren egoera egonkorreko banaketa kalkulatuko da $\mu = \infty$ kasuan eta 6.4.2. Sekzioan $\mu < \infty$ kasuan.

6.4.1 Zerbitzu tasa infinitua

Zerbitzu tasa infinitua den kasuan sistemak ilara husten du zerbitzaria aktibatzeko erabakia hartu bezain laster. Kasu honetan, n atari-politikapean, egoera m da non $m \in E = \{0, 1, \ldots, n-1\}$. $m = n-1$ egoeran iritsiera bat gertatzen den bezain laster sistema berehala husten da.

Izan bedi π_m $m \in \{0, \ldots, n-1\}$ egoerako egoera-egonkorreko probabilitatea. Beraz, hurrengo oreka ekuazioak betetzen dira:

$$\lambda \pi_{m-1} = \theta m \pi_m + \lambda \pi_{n-1}, \quad \forall\, 0 < m \leq n-1,$$

eta normalizazio ekuazioa $\sum_{i=0}^{n-1} \pi_i = 1$. Orain oreka ekuazioak askatuko dira, hau da,

$$\begin{aligned}
\pi_m &= (m+1)\frac{\theta}{\lambda}\pi_{m+1} + \pi_{n-1} \\
&= \pi_{n-1} + (m+1)\frac{\theta}{\lambda}\left((m+2)\frac{\theta}{\lambda}\pi_{m+2} + \pi_{n-1}\right) \\
&= \pi_{n-1}\left(1 + \frac{\theta}{\lambda}(m+1)\right) + \pi_{m+2}\left(\frac{\theta}{\lambda}\right)^2 (m+1)(m+2) \\
&= \pi_{n-1}\left(1 + \frac{\theta}{\lambda}(m+1) + \left(\frac{\theta}{\lambda}\right)^2 (m+1)(m+2)\right) \\
&\quad + \pi_{m+3}\left(\frac{\theta}{\lambda}\right)^3 (m+1)(m+2)(m+3) \\
&= \ldots \\
&= \pi_{n-1}\left(1 + \sum_{i=1}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i \frac{(m+i)!}{m!}\right).
\end{aligned}$$

Beraz,

$$\pi_m = \pi_{n-1}\left(\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i \frac{(m+i)!}{m!}\right),$$

edozein $m = 0, 1, \ldots, n-1$. Gainera, normalizazio ekuaziotik, hurrengoa lortu daiteke

$$\pi_{n-1} = \left(\sum_{m=0}^{n-1}\left(\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i \frac{(i+m)!}{m!}\right)\right)^{-1}. \tag{6.4.1}$$

Beraz, π_m-ren espresioa lortu da edozein $0 \leq m \leq n-1$ bada.

6.4.2 Zerbitzu tasa finitua

Zerbitzu tasa finitua den kasuan, sistemaren egoera (m, a) da non m ilaran zain dauden bezero kopurua den eta $a \in \{0, 1\}$-k zerbitzaria lanean ($a = 1$) edo libre ($a = 0$) dagoen adierazten duen. Ohartu n ataripean, zerbitzaria libre dagoen bitartean, ilaran dauden bezero kopuruak *i.e.*, m-k, balioak $\{0, \ldots, n-1\}$ multzoan hartzen ditu. Azken honek adierazten du, iritsiera bat gertatzen den unean $m = n-1$ egoeran, zerbitzariak zerbitzua aktibatzen duela n bezeroko sorta batekin. Zerbitzaria aktiboa den bitartean, bezero sorta prozesatua den bitarteko denbora (zeinak banaketa esponentziala duen $1/\mu$ batez bestekoa duena), ilaran zain dauden bezero kopuruak $m \in \mathbb{N} \cup \{0\}$ betetzen du. Hau da, n ataria iritsi den arren, zerbitzaria aktibatzea atzeratu egiten da zerbitzariak aurreko bezero sorta prozesatzen bukatu duen arte.

Izan bedi $\pi(m,a)$ (m,a) egoeraren egoera-egonkorreko probabilitatea, edozein $m \geq 0$ denean, eta $a \in \{0,1\}$, eta izan bedi $\pi(m,0) = 0$ edozein $m \geq n$. Aurreko sekzioan bezela, n-rekiko dependentzia alde batera utzi da notaziotik. Orduan, $\pi(m,a)$ edozein $m \in \mathbb{N} \cup \{0\}$ eta $a \in \{0,1\}$ badira hurrengo oreka ekuazioetatik ondoriozta daiteke edozein $m \in \mathbb{N}$ bada:

$$(\lambda + m\theta + \mu)\pi(m,1) = \lambda\pi(m-1,1) + (m+1)\theta\pi(m+1,1), \tag{6.4.2}$$

eta edozein $0 \leq m \leq n-1$ bada

$$(\lambda + m\theta)\pi(m,0) = \lambda\pi(m-1,0) + \mu\pi(m,1) + (m+1)\theta\pi(m+1,0). \tag{6.4.3}$$

(6.4.2). eta (6.4.3) ekuazioek definitutako oreka ekuazioak ebatzi ahal izateko eurei dagokien funtzio sortzaileak erabiliko dira. Ohartu (6.4.3). Ekuazioan egoera-egonkorreko probabilitateak zerbitzaria libre deneko denboraldian, zerbitzaria lanpetuta dagoeneko denboraldiko egoera-egonkorreko probabilitateen menpekoa da. Beraz, $\pi(m,1)$-ren espresio esplizitua aurkituko da edozein m-rentzat, eta espresio hauek erabiliz zerbitzaria libre deneko probabilitateak lortuko dira, *i.e.*, $\pi(m,0)$ edozein $n-1 \geq m \geq 0$ bada. Probabilitateen espresio esplizituak 6.2. Proposizioan aurki daitezke. Espresio hauek lortu ahal izateko aurrera eraman diren kalkuluak 6.7.2. Eranskinean aurki daitezke, eta zerbitzaria lanpetuta dagoeneko denboraldikoak 6.7.2. Eranskinean.

6.2 Proposizioa. *Izan bedi* $a_1(0) := 1$,

$$a_1(1) := \frac{\lambda+\mu}{\theta} - \frac{e^{\lambda/\theta}}{\sum_{i=0}^{\infty} \frac{(\lambda/\theta)^i}{i!(\mu/\theta+i)}}, \text{ and,}$$

$$a_1(m) := \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}\left(\frac{\sum_{j=0}^{\infty}\frac{\left(\frac{\lambda}{\theta}\right)^j(-\ell_{k-1}(-\frac{\mu}{\theta}-j+1))}{j!}}{\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)}}\sum_{i=0}^{m-k}\binom{m-k}{i}\left(\frac{\lambda}{\theta}\right)^{m-k-i}\ell_i\left(\frac{\mu}{\theta}\right)\right),$$

edozein $m \geq 2$, *non* $\ell_k(x)$ *Pochhammer funtzioa den edo rising factorial funtzioa. Izan bedi* $a_0^n(n-1) := \frac{\mu}{\lambda}\sum_{m=0}^{n-1} a_1(m)$,

$$a_0^n(m) := \left(\frac{\mu}{\lambda}\sum_{r=0}^{n-1}a_1(r)\right)\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!} - \frac{\mu}{\lambda}\sum_{r=m+1}^{n-1}a_1(r)\left(\sum_{i=0}^{r-m-1}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\right),$$

edozein $n-2 \geq m > 0$, *eta* $a_0^n(0) := \frac{\mu}{\lambda} + \frac{\theta}{\lambda}a_0^n(1)$. *Orduan*

$$\pi(0,1) = \left(\sum_{m=0}^{n-1}a_0^n(m) + \sum_{m=0}^{\infty}a_1(m)\right)^{-1},$$

denez $\pi(m,0) = a_0^n(m)\pi(0,1)$, *eta* $\pi(m,1) = a_1(m)\pi(0,1)$, *zeinetatik* $\sum_{m=0}^{n-1}\pi(m,0) + \sum_{m=0}^{\infty}\pi(m,1) = 1$ *ondorioztatzen den, eta* $\pi(m,0)$ *eta* $\pi(m,1)$ (6.4.2) *eta* (6.4.3) *ekuazioen emaitzak diren.*

$\pi(m,0)$-ren espresioa lortu delarik edozein $n-1 \geq m \geq 0$ denean, eta $\pi(m,1)$ edozein $m \geq 0$ denean, sisteman dauden batezbesteko bezero kopurua kalkulatu daiteke n politikapean, baita zerbitzaria lanpetuta dagoeneko batezbesteko denbora. Hurrengo sekzioan ikusiko da honek atari politiken karakterizazio bat ahalbideratzen duela.

6.5 Atari-politika optimoaren karakterizazioa

Sekzio honetan atari-politika optimoa karakterizatuko da goian kalkulatu diren egoera-egonkorreko probabilitateak erabiliz. Karakterizazio honek, c_s^∞ eta c_s kostuekiko dependentzia du. Zerbitzariaren abiadura infinitua den kasuan ($\mu = \infty$), atari-politikaren karakterizazioa esplizitua da, eta soluzioa 6.3. Proposizioan aurkeztuko da. Zerbitzu abiadura finituko kasuan ($\mu < \infty$), politika optimoaren karakterizazioa funtzio baten optimizazioan oinarritzen da, ikusi 6.4. Proposizioa.

6.1. Proposizioan P problemarentzat atari politikak optimoak direla ikusi da, orduan P problema

$$\min_{n\in\{0,1,\dots\}} \left(\tilde{c}\mathbb{E}(N^n) + c_s^\infty \lambda \mathbb{E}(\mathbf{1}_{\{N^n=n-1\}})\right),$$

idatz daiteke $\mu = \infty$ kasuan, non $\mathbb{E}(\mathbf{1}_{\{N^n=n-1\}}) = \pi_{n-1}^n$ eta

$$\min_{n\in\{0,1,\dots\}} \left(\tilde{c}\mathbb{E}(N^n) + c_s \mathbb{E}(\mathbf{1}_{\{S^n(N^n)=1\}})\right),$$

$\mu < \infty$ kasuan, non $\mathbb{E}(\mathbf{1}_{\{S^n(N^n)=1\}}) = \sum_{m=0}^\infty \pi^n(m,1)$. Orain hurrengo notazioa definituko da, $P_b^n = \pi_{n-1}^n$ $\mu = \infty$ kasuan eta $P_b^n = \sum_{m=0}^\infty \pi^n(m,1)$ $\mu < \infty$ kasuan. Hemendik aurrera $c_s = c_s^\infty \lambda$ erabiliko da $\mu = \infty$ kasuan.

Hurrengo proposizioan atari politika optimoaren errepresentazio bat proposatuko da.

6.3 Proposizioa. *Izan bedi $\alpha(n)$*

$$\alpha(n) := \tilde{c}\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-1})}{P_b^{n-1} - P_b^n},$$

edozein $n > 1$ bada. $\alpha(n)$ ez-beherakorra bada n-n, P_b^n ez-beherakorra da n-n, eta $\alpha(n) \leq c_s < \alpha(n+1)$, bada orduan n da P problemaren politika optimoa.

Hurrengo leman 6.3. Proposizioko baldintzak betetzen direla frogatuko da $\mu = \infty$ kasuan. Frogapena 6.7.4. Eranskinean aurki daiteke.

6.1 Lema. *Izan bedi $\mu = \infty$, eta π_m^n 6.4.1. Sekzioan definitu dena, n indizea gehitu ondoren. Orduan*

- *$P_b^n = \pi_{n-1}^n$ ganbila eta ez-gorakorra da.*
- *$\alpha(n)$ funtzioa, 6.3. Proposizioan definitu dena, ez-beherakorra da.*

6.1 Korolarioa. *Izan bedi $\mu = \infty$, $c_s = c_s^\infty \lambda$ eta definitu $\alpha(1) := -\infty$. Orduan, $\alpha(n) \leq c_s < \alpha(n+1)$ bada edozein $n \geq 1$, n P problemarentzat atari-politika optimoa da.*

Frogapena. Frogapena 6.3. Proposiziotik eta 6.1. Lematik ondorioztatzen da. □

$\mu < \infty$ den kasuan ezin izan da α ez-beherakorra dela frogatu. Atari optimoak beraz, modu ezberdin batean karakterizatu behar da. Karakterizazio hau hurrengo proposizioan aurki daiteke.

6.4 Proposizioa. *Izan bedi* $\mathcal{N}_i = \mathbb{N}\backslash\{0,\ldots,n_i\}$ n_i *batetarako, eta demagun* P_b^n *ez-gorakorra dela. Definitu* $\beta(\cdot)$ *hurrengo moduan: i. pausua Kalkulatu*

$$\beta(n_i) := \inf_{n\in\mathcal{N}_{i-1}} \tilde{c}\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n_{i-1}})}{P_b^{n_{i-1}} - P_b^n}, i \geq 1, \tag{6.5.1}$$

eta deitu n_i $n \in \mathcal{N}_{i-1}$ *minimizatzaile handienari non* (6.5.1) *minimizatzen den.* $n_i = \infty$ *bada gelditu, bestela egin jauzi* $i+1$*. pausura.*

Orduan, $\beta(n_i)$ *ez-beherakorra da* n_i*-n eta* $\beta(n_i) \leq c_s < \beta(n_i+1)$ *bada, orduan* n_i *optimoa da* P *problemarentzat.*

Ez hori bakarrik, $c_s < \beta(n_1)$ *bada, orduan soluzio optimoa beti zerbitzua aktibatzea da.*

Frogapena. Lehenik eta behin $\beta(n_i)$ n_i-n ez-beherakorra dela frogatuko da. Gogoratu definizioz

$$\frac{\mathbb{E}(N^{n_i}) - \mathbb{E}(N^{n_{i-1}})}{P_b^{n_{i-1}} - P_b^{n_i}} \leq \frac{\mathbb{E}(N^{n_{i+1}}) - \mathbb{E}(N^{n_{i-1}})}{P_b^{n_{i-1}} - P_b^{n_{i+1}}},$$

dela eta orduan $(\mathbb{E}(N^{n_i})-\mathbb{E}(N^{n_{i-1}}))(P_b^{n_{i-1}}-P_b^{n_{i+1}}) \leq (\mathbb{E}(N^{n_{i+1}})-\mathbb{E}(N^{n_{i-1}})))(P_b^{n_{i-1}}-P_b^{n_i})$. $\mathbb{E}(N^{n_i})(P_b^{n_{i-1}}-P_b^{n_i})$ gaia batu eta kenduz inekuazioaren eskuinaldean, eta hainbat kalkulu egin ostean $\beta(n_i) \leq \beta(n_i+1)$ lortzen da.

$\beta(\cdot)$ ez-beherakorra dela frogatuz, atari-politika optimoa n_i $\beta(n_i) \leq c_s < \beta(n_i+1)$ bada 6.3. Proposizioan frogatu den antzeko moduan froga daiteke. □

Hurrengo lemak 6.4. Proposizioko karakterizazioa betetzen dela bermatzen du $\mu < \infty$ kasuan. Frogapena 6.7.5. Eranskinean aurki daiteke.

6.2 Lema. *Demagun* $\mu < \infty$ *eta* $\pi^n(m,a)$ *6.4.2. Sekzioan definitu den bezela, edozein* $m \geq 0$ *eta* $a \in \{0,1\}$*,* n *indizea gehitu ondoren. Orduan* $P_b^n = \pi^n(0,1)\sum_{m=0}^{\infty} a_1(m)$*, ez-gorakorra da.*

Soluzio optimoa karakterizatu da $\beta(n_i)$ funtzioaren bitartez, zeina 6.4. Proposizioan definitu den, hurrengo koroloarioan $\mu < \infty$ kasuan.

6.2 Korolarioa. *Izan bedi* $\mu < \infty$ *eta* $\beta(n_i) \leq c_s < \beta(n_i+1)$ *non* β *6.4. Proposizioan definitu den. Orduan* n_i P *problemarentzat atari-politika optimoa da.* $c_s < \beta(n_1)$ *bada, orduan* 0 *da atari-politika optimoa (beti zerbitzatzea).*

Frogapena. Frogapena 6.4. Proposiziotik eta 6.2. Lematik ondorioztatzen da. □

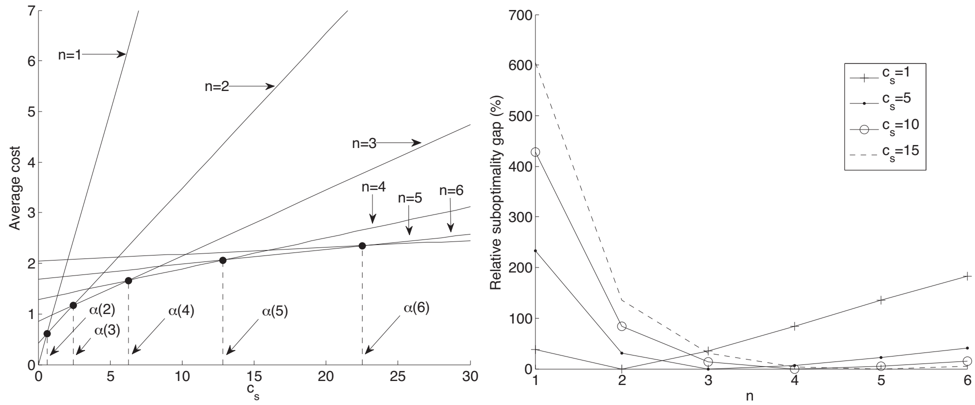
Hurrengo sekzioan atari-politika optimoa aztertuko da numerikoki bi kasuetarako $\mu = \infty$ eta $\mu < \infty$.

6.6 Adibideak

Sekzio honetan atari politika optimoen propietateak aztertuko dira, zeinak 6.5. Sekzioan karakterizatu diren, adibide ezberdinen bitartez. 1 eta 2 adibideetan atari politika optimoa irudikatu da c_s^{∞} aktibatze kostu ezberdinetarako eta c_s zerbitzu kostu ezberdinetarako, bi kasuetan $\mu = \infty$ eta $\mu < \infty$, hurrenez hurren, optimoak ez diren atari-politiken errendimendua aztertu da. 3 eta 4 adibideetan atari politikak θ parametroaren arabera nola aldatzen diren aztertu da $\mu = \infty$ kasuan eta θ eta μ parametroen arabera $\mu < \infty$ kasuan.

Taula 6.1: 1 Adibidea: c_s aktibatze kostu minimoa n optimoa izan dadin.

n	1	2	3	4	5	6
c_s	$-\infty$	0.6096	2.4359	6.2595	12.8192	22.5343



Irudia 6.2: Ezkerraldean batez besteko kostua irudikatu da atari-politika ezberdinentzat eta c_s aldatzen denean. Eskuinaldean errore erlatiboa irudikatu da atari-politika ezberdinentzak atari-politika optimoarekiko.

1 Adibidea: Izan bedi $\lambda = 4$, $\mu = \infty$, $\theta = 1.5$ eta $\tilde{c} = 1$. Orduan c_s-ren balio minimioa, non $c_s = c_s^\infty \lambda$, n optimoa izan dadin P problemarentzat, $\alpha(n)$ funtzioak definitzen du, zeinaren balioak 6.1. Taulan aurkeztu diren. 6.2. Irudian (ezkerraldean) soluzio optimoa irudikatu da, zeinetarako n politikapeko $\mathbb{E}(N^n) + c_s P_b^n$ batezbesteko kostua irudikatu den, c_s-ren balio ezberdinetarako. $n = 6$-rainoko emaitzak aurkeztuko dira, eta ohartu karakterizatio bat aurki daitekela edozein n-rentzat. 6.2. Irudian (eskuinaldean) optimoak ez diren errore erlatiboa aurkeztu da soluzio optimoarekiko konparatuz. Hauen errendimendua oso kaxkarra da.

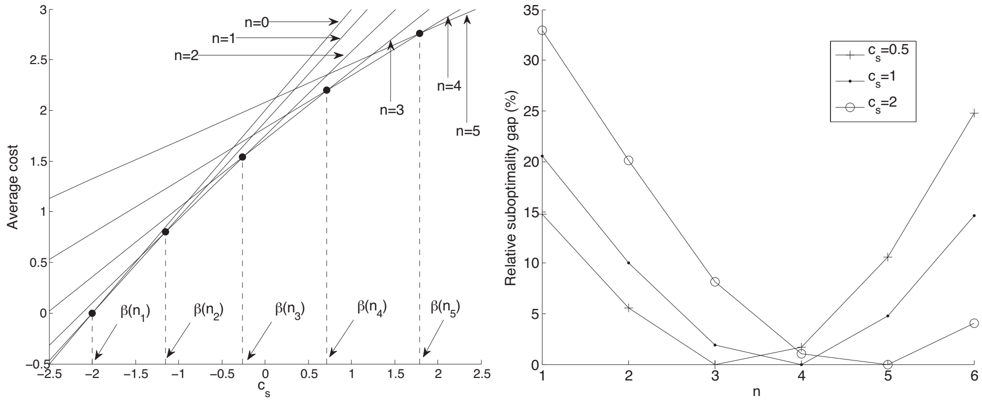
2 Adibidea: Izan bedi $\lambda = 2$, $\mu = 0.5$, $\theta = 0.5$ eta $\tilde{c} = 1$. Orduan c_s-ren balio minimoa, n P problemaren soluzio optimoa izan dadin $\beta(n_i)$ balioek zehazten dute, zeina $\alpha(i)$-rekin bat datorren kasu honetan. $\beta(n_i)$-ren balioak 6.2. Taulan aurkeztu dira $n_i = i$ denean eta $i = 5$ arte. Ohartu $c_s > 0$ dela onartuz $n = 0, 1, 2$ politikak ez direla inoiz optimoak kasu partikular honetan, zeinak adierazten duen 1 edo 2 bezero badaude zerbitzatzeko zain zerbitzariak ez dituela zerbitzuan hartuko.

Emaitza hauek 6.3. Irudian (ezkerraldean) aurkeztu dira, zeinetan $\mathbb{E}(N^n) + c_s P_b^n$ batez besteko kostua irudikatu den n politikapean c_s-ren balio aldakorretarako. 6.3. Irudian (eskuinaldean) errore erlatiboa irudikatu da atari-politika ezberdinentzat atari-politika optimoarekiko, eta ohartu politika hauek errendimendu kaxkarra erakusten dutela.

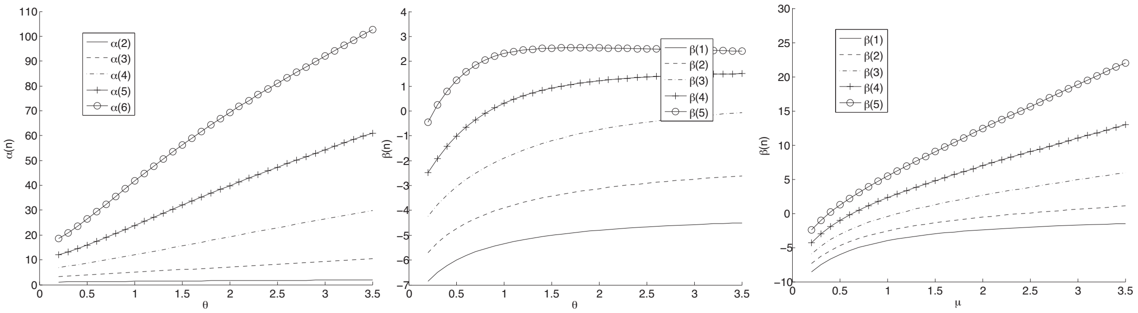
Hurrengo bi adibideetan θ eta μ aldatzen diren heinean politiken aldaketa aztertzen da.

Taula 6.2: 2 Adibidea: c_s aktibazio kostu minimoa n optimoa izan dadin

n	0	1	2	3	4	5
c_s	$-\infty$	-2	-1.151	-0.2581	0.7157	1.7937



Irudia 6.3: Ezkerraldean batez besteko kostua irudikatu da atari-politika ezberdinentzat eta c_s aldatzen denean. Eskuinaldean errore erlatiboa irudikatu da atari-politika ezberdinentzak atari-politika optimoa-rekiko.



Irudia 6.4: α eta β-ren balioak irudikatu dira θ eta μ aldatzen diren heinean. Ezkerraldean $\mu = \infty$ kasua aztertu da θ aldatzen den heinean. Erdian eta eskuinaldean $\mu < \infty$ kontsideratu da θ eta μ aldatzen diren heinean, hurrenez hurren.

3 Adibidea: Izan bedi $\lambda = 4$, $\mu = \infty$ eta $c = \delta = 1$ eta θ $[0, 3.5]$ tartean aldatzen da. Ohartu 6.4. Irudian (ezkerraldean) c_s-ren balio minimioa, non $c_s = c_s^\infty \lambda$ den, zeinetarako n atari-politika optimoa den gorakorra da θ-rekiko. Honek esan nahi du aktibazio kostua finkoa denean eta uzte tasak gorakorrak direnean, sistemak zerbitzua lehenago aktibatzea erabakitzen duela uzteen penalizazioak ekiditeko.

4 Adibidea: Izan bedi $\lambda = 2$ eta $c = \delta = 1$. Lehenik eta behin $\mu = 0.5$ kontsideratu da eta θ aldakorra dela kontsideratu $[0, 3.5]$ tartean, ikusi 6.4. Irudia (erdialdean). Ohartu θ handitzen den heinean, $\beta(n)$ konstantea bilakatzen dela. Fenomeno hau hurrengoak azaltzen du: bezeroen uzteak lanpetutako denboraldian zein denboraldi librean gerta daitezke eta beraz, zerbitzu kostua finkoa bada atari politika mantentzen da θ hazten den heinean. Bigarren, $\theta = 0.5$ dela onartuko da eta μ aldakorra dela $[0, 3.5]$ tar-tean. Ohartu 6.4. Irudian (eskuinaldean), zerbitzua zenbat eta azkarragoa izan, orduan eta atari politika txikiagoa dela optimoa c_s finko mantenduz.

6.7 Eranskina

6.7.1 6.1. Proposizioaren frogapena: $\mu < \infty$ kasua

Lehenik eta behin SRT aplikatu ahal izateko baldintzak betetzen direla frogatuko da. Beranduago atari politikak optimoak direla ikusiko da $\mu < \infty$ kasuan.

[25, 3.1. Teorema]-ko baldintzen egiaztapena

Lehenik eta behin frogatuko da [25, 3.1. Teorema] artikuloan aipatzen diren baldintzak P problemak betetzen dituela. Izan bedi $E = \mathbb{N} \cup \{0\}$, eta definitu $h(m) = \mathrm{e}^{\epsilon m}$, orduan, 3.1. Definizioaz, h momentu funtzio at dela lortzen da. Bete beharreko baldintzen enuntziatua 3.7.1. Eranskinean aurkitzen da, non baldintza hauek egiaztatu diren uzteak gerta daitezkeen ilara klase anitz batentzat.

Izan bedi $\mu < \infty$ eta $k_1 = (k_{11}, k_{12})$, orduan lehenengo baldintza 3.7.1. Eranskinean aipatu den legez, hurrengoa da: aurkitu $\epsilon > 0$ eta $\tilde{M} > 0$ non edozein $m \geq \tilde{M}$ bada

$$\begin{aligned}
&\lambda\left(1 - \frac{m}{L}\right)\mathrm{e}^{\epsilon(m+1)}(1 - S^{\phi}(m)) + \theta m \mathrm{e}^{\epsilon(m-1)}(1 - S^{\phi}(m)) + \left(\lambda\left(1 - \frac{m}{L}\right) + \theta m\right)S^{\phi}(m) \\
&- \left(\lambda\left(1 - \frac{m}{L}\right) + \theta m\right)\mathrm{e}^{\epsilon m} \leq -k_{11}\mathrm{e}^{\epsilon m}, \\
&\mu \mathrm{e}^{\epsilon m} + \lambda\left(1 - \frac{m}{L}\right)\mathrm{e}^{\epsilon(m+1)} + \theta m \mathrm{e}^{\epsilon(m-1)} - \left(\lambda\left(1 - \frac{m}{L}\right) + \mu + \theta m\right)\mathrm{e}^{\epsilon m} \leq -k_{12}\mathrm{e}^{\epsilon m},
\end{aligned}$$

eta k_{11} eta k_{12} konstanteak direlarik. Azken ekuazioan, lehen inekuazioa $(m, 0)$ egoerari dagokio eta bigarren inekuazioa $(m, 1)$ egoerari. Hainbat kalkulu egin ostean azken bi inekuazioak

$$\lambda\left(1 - \frac{m}{L}\right)(\mathrm{e}^{\epsilon} - 1 + S^{\phi}(m)(\mathrm{e}^{-\epsilon m} - \mathrm{e}^{\epsilon})) + \theta m(\mathrm{e}^{-\epsilon} - 1 + S^{\phi}(m)(\mathrm{e}^{-\epsilon m} - \mathrm{e}^{-\epsilon})) \leq -k_{11},$$

non $\lambda\left(1 - \frac{m}{L}\right)\left(\mathrm{e}^{\epsilon} - 1 + S^{\phi}(m)(\mathrm{e}^{-\epsilon m} - \mathrm{e}^{\epsilon})\right)$ goitik bornatua den konstante batez, κ izendatuko dena, eta $\theta m(\mathrm{e}^{-\epsilon} - 1 + S^{\phi}(m)(\mathrm{e}^{-\epsilon m} - \mathrm{e}^{-\epsilon})) \leq 0$, beraz, aurki daiteke $\tilde{M}$ nahiko handia $\theta m(\mathrm{e}^{-\epsilon} - 1 + S^{\phi}(m)(\mathrm{e}^{-\epsilon m} - \mathrm{e}^{-\epsilon})) \leq -\kappa$ izan dadin eta

$$\lambda\left(1 - \frac{m}{L}\right)(\mathrm{e}^{\epsilon} - 1) + \theta m(\mathrm{e}^{-\epsilon} - 1) \leq -k_{12},$$

idatz daitezke, non $\lambda\left(1 - \frac{m}{L}\right)(\mathrm{e}^{\epsilon} - 1)$ goitik bornatua den eta nahi bezain negatiboa egin daitekeen $\tilde{M}$ handia bada.

Egiaztatu beharreko bigarren baldintza, 3.7.1. Eranskinean aipatu den legez, $q^{\phi,L}(\cdot,\cdot)$ jarraia izatea da $S^{\phi}(N^{\phi}(t))$-n eta L-n, zeina definizioz betetzen den.

Atari-politika optimoak

Izan bedi $V(m, a)$ P problemari dagokion balio funtzioa $\mu < \infty$ kasuan, eta izan bedi g politika optimo batek eragindako batez besteko kostua. Gogoratu sistemaren egoera (m, a) dela, non m ilaran zain dauden bezeroen kopurua den eta $a \in \{0, 1\}$-k zerbitzaria lanpetua ($a = 1$) edo libre dagoen ($a = 0$) adierazten duen. $V(m, a)$ balio funtzioa edozein $m \geq 0$ eta $a \in \{0, 1\}$ denean Bellman ekuazioa betetzen du, ikusi

A.1. Eranskina, hau da,

$$\begin{aligned}(\lambda+\mu+\theta m)V(m,0)+g=&\tilde{c}m+\min\{\lambda V(m+1,0)+\theta mV((m-1)^+,0)\\&+\mu V(m,0),c_s+\lambda V(1,1)+\theta mV(0,1)+\mu V(0,0)\},\end{aligned}$$

eta

$$(\lambda+\mu+\theta m)V(m,1)+g=\tilde{c}m+c_s+\lambda V(m+1,1)+m\theta V((m-1)^+,1)+\mu V(m,0).$$

Orain frogatuko da Bellman ekuazioa ebazten duen politika optimoa atari-politika bat dela, hau da, aktibo akzioa optimoa bada m egoeran orduan aktibo akzioa optimoa da $m'\geq m$ egoeran ere. Hori frogatu ahal izateko, definitu

$$\begin{aligned}f(m,0):=&\tilde{c}m+\lambda V(m+1,0)+\mu V(m,0)+\theta mV(m-1,0),\\f(m,1):=&\tilde{c}m+c_s+\lambda V(1,1)+\theta mV(0,1)+\mu V(0,0),\end{aligned}$$

eta $\varphi(m)=\min\left(b\in\arg\min_{a\in\{0,1\}}f(m,a)\right)$. Nahikoa da orduan $\varphi(m')\geq\varphi(m)$ dela ikustea edozein $m'\geq m$ bada. Izan bedi $a\geq\varphi(m')$. $\varphi(\cdot)$-ren definizioz

$$f(m',\varphi(m'))-f(m',a)\leq 0. \tag{6.7.1}$$

Froga dezagun orain $V(m,0)$ *subadditive* dela [76], hau da, edozein $m'\geq m$ bada eta $a\in\{0,1\}$

$$f(m',a)+f(m,\varphi(m'))\leq f(m',\varphi(m'))+f(m,a). \tag{6.7.2}$$

Lehenik eta behin onartu $\varphi(m')=0$ dela eta $a=0$, orduan (6.7.2) berehalakoa da, antzeko moduan (6.7.2) betetzen da $\varphi(m')=1$ bada eta $a=1$. $\varphi(m')=0$ eta $a=1$ kasua falta da aztertzea, zeinarentzat (6.7.2) ekuazioa $f(m,0)-f(m,1)\leq f(m',0)-f(m',1)$, den. P problema $\mu<\infty$ kasuan ez denez uniformizagarria, SRT metodoa erabiliko da. L parametroarekin trunkatuko da egoera espazioa, eta iritsiera tasak moldatuko dira *i.e.*, $q(m,m+1)=\lambda\left(1-\frac{m}{L}\right)^+$ edozein $m\leq L$ bada. $V^L(\cdot,\cdot)$ bidez izendatuko da trunkatutako sistemaren balio funtzioa. SRT aplikatu aha izateko beharrezko baldintzak [25, 3.1. Teorema] betetzen direnez, $V^L\to V$ da $L\to\infty$ den heinean eta balio funtzioak egitura propietateak mantentzen ditu. Onartu orokortasunik galdu gabe, $\lambda+\mu+\theta L=1$. Ordain, *value iteration* algoritmoa aplikatuz, ikusi 1.3.3. Sekzioa, sistema trunkaturi,

$$\begin{aligned}V_{t+1}^L(m,0)=\tilde{c}m+\min\{&\lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0)+\left(\lambda\frac{m}{L}+\mu+\theta(L-m)\right)V_t^L(m,0)+\theta mV_t^L(m-1,0),\\&c_s+\lambda V_t^L(1,1)+\theta LV_t^L(0,1)+\mu V_t^L(0,0)\},\end{aligned} \tag{6.7.3}$$

lortzen da, eta

$$\begin{aligned}&V_{t+1}^L(m,1)=\tilde{c}m+c_s+\lambda\left(1-\frac{m}{N}\right)V_t^L(m+1,1)+\mu V_t^L(m,0)\\&+\left(\lambda\frac{m}{N}+(N-m)\theta\right)V_t^L(m,1)+m\theta V_t^L(m-1,1),\end{aligned}$$

$V^L_{t+1}(m,a) - V^L_t(m,a) = g$ denez. Orduan nahikoa da hurrengoa frogatzea $f^L_t(m,0) - f^L_t(m,1) \leq f^L_t(m',0) - f^L_t(m',1)$, non

$$\begin{aligned}f^L_t(m,0) =&\lambda \left(1 - \frac{m}{L}\right) V^L_t(m+1,0) + \theta m V^L_t(m-1,0) + \left(\lambda \frac{m}{L} + \mu + \theta(L-m)\right) V^L_t(m,0),\\ f^L_t(m,1) =&\tilde{c}m + c_s + \lambda V^L_t(1,1) + \theta L V^L_t(0,1) + \mu V^L_t(0,0).\end{aligned}$$

V^L funtzio ez-beherakorra da eta $f^L_t(m,0) - f^L_t(m,1) \leq f^L_t(m',0) - f^L_t(m',1)$ edozein $m' \geq m$ bada eta edozein t, bien frogapena hemen behean aurki daiteke. *Value iteration* argumentua erabiliz $f^L_t \to f^L$ $t \to \infty$ den heinean, eta SRT erabiliz $V^L \to V$ lortzen da eta $f^L \to f$ puntuz puntu. Orduan, $V^L(\cdot,0)$ *subadditive* izateak $V(\cdot,0)$ *subadditive* izatea inplikatzen du. (6.7.1) eta (6.7.2) betetzen direla frogatuz, eta biak konbinatuz, hurrengoa lortzen da edozein $a \geq \varphi(m')$ eta $m' \geq m$ badira

$$\begin{aligned}f(m,\varphi(m')) &\leq f(m',\varphi(m')) - f(m',a) + f(m,a)\\ &\leq f(m,a).\end{aligned}$$

Orduan $\varphi(m) \leq \varphi(m')$, zeinak frogapena amaitzen duen.

$V^L(\cdot,0)$ **ez-beherakorra izatearen frogapena.** $V^L(\cdot,0)$ funtzio ez-beherakorra dela frogatzeko $V^L_0(m) = 0$ definituko da edozein $m \leq L$ bada. g batez besteko kostu optimoa bada, eta orokortasunik galdu gabe $\lambda + \mu + \theta L = 1$ onartzen bada Bellman-en ekuazioa

$$\begin{aligned}V^L_{t+1}(m,0) = \tilde{c}m + \min\Bigg(&\lambda \left(1 - \frac{m}{L}\right) V^L_t(m+1,0)\\ &+ \left(\lambda \frac{m}{L} + \mu + \theta(L-m)\right) V^L_t(m,0) + \theta m V^L_t(m-1,0);\\ &c_s + \lambda V^L_t(1,1) + \theta L V^L_t(0,1) + \mu V^L_t(0,0)\Bigg),\end{aligned} \tag{6.7.4}$$

eta

$$\begin{aligned}V^L_{t+1}(m,1) =&\tilde{c}m + \lambda \left(1 - \frac{m}{L}\right) V^L_t(m+1,1) + \mu V^L_t(m,0)\\ &+ \left(\lambda \frac{m}{L} + (L-m)\theta\right) V^L_t(m,1) + m\theta V^L_t(m-1,1),\end{aligned} \tag{6.7.5}$$

da, izan ere $V^L_{t+1}(m,0) - V^L_t(m,0) = g$.

Lehenik eta behin $V^L_t(m,0)$ m-n ez-beherakorra dela frogatuko da edozein t bada. Indukzioz: $V^L_0(m,0) \geq V^L_0(m',0)$ edozein $m \geq m' \geq 0$ dela frogatuko da zeinak $V^L_1(m,0) \geq V^L_1(m',0)$ inplikatzen duen, eta beranduago $V^L_t(m,a)$ m-n ez-beherakorra izateak, $V^L_{t+1}(m,1)$ ez-beherakorra izatea inplikatzen duela frogatuko da. Definizioz $V^L(m,a)$ m egoeran hasi izanaren kostu totalaren diferentzia asintotikoa da erreferentzi den egoera batean hasi ordez. Orokortasunik galdu gabe, erreferentzi egoera 0 dela onartuko da. $V^L_0(m,a) = 0$ definituko da edozein $m \geq 0$ eta $a \in \{0,1\}$ badira, orduan (6.7.4). eta (6.7.5). Ekuaziotik $V^L_1(m,a) = \tilde{c}m$ lortuko da. $\tilde{c} > 0$ eta $V^L_1(m,a) \geq V^L_1(m-1,a)$ edozein $m \geq 1$-tarako baita. Orain onartu $V^L_t(m,0)$ ez-beherakorra dela, eta frogatu $V^L_{t+1}(m,0) \geq V^L_{t+1}(m-1,0)$ edozein $L \geq m \geq 1$, hau da, (6.7.4). ekuazioa ordezkatuz, hurrengoa frogatzearen baliokidea da:

$$\tilde{c}m + \min\Bigg(\lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0) + \theta m V_t^L(m-1,0) + \left(\lambda\frac{m}{L} + \mu + \theta(L-m)\right)V_t^L(m,0), c_s + \lambda V_t^L(1,1)$$
$$+ \theta N V_t^L(0,1) + \mu V_t^L(0,0)\Bigg)$$
$$\geq \tilde{c}(m-1) + \min\Bigg(\lambda\left(1-\frac{m-1}{L}\right)V_t^L(m,0) + \left(\lambda\frac{m-1}{L} + \mu + \theta(L-m+1)\right)V_t^L(m-1,0)$$
$$+ \theta(m-1)V_t^L((m-2)^+,0), c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0)\Bigg). \tag{6.7.6}$$

Orain (6.7.6). Inekuazioa betetzen dela ikusiko da edozein m eta $m-1$ egoeretako akzio konbinaziotarako. Lehenik eta behin bai m eta bai $m-1$ egoeretan akzio pasiboa dela optimoa onartuko da, orduan (6.7.6). Ekuazioa

$$\tilde{c}m + \lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0) + \theta m V_t^L(m-1,0) + \left(\lambda\frac{m}{L} + \mu + \theta(L-m)\right)V_t^L(m,0)$$
$$\geq \tilde{c}(m-1) + \lambda\left(1-\frac{m-1}{L}\right)V_t^L(m,0) + \left(\lambda\frac{m-1}{L} + \mu + \theta(L-m+1)\right)V_t^L(m-1,0)$$
$$+ \theta(m-1)V_t^L((m-2)^+,0),$$

idatz daiteke, zeinak hainbat kalkuluren ostean

$$\tilde{c} + \lambda\left(1-\frac{m}{L}\right)\Delta V_t^L(m+1,0) + \theta(m-1)\Delta V_t^L(m-1,0) + \left(\lambda\frac{m-1}{L} + \mu + \theta(L-m)\right)\Delta V_t^L(m,0) \geq 0,$$

idatz daitekeen edozein $L \geq m \geq 1$ eta $\Delta V_t^L(m,0) = V_t^L(m,0) - V_t^L((m-1)^+,0)$ badira. $\Delta V_t^L(m,0) \geq 0$ denez edozein $L \geq m \geq 0$-tarako, azken inekuazio hau betetzen da. Orain(6.7.6) frogatuko da aktibo akzioa denean optimoa m eta $m-1$ egoeretan, orduan, (6.7.6). Inekuazioa

$$\tilde{c}m + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0)$$
$$\geq \tilde{c}(m-1) + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0),$$

idatz daiteke, zeina $\tilde{c} \geq 0$ idatz daitekeen, eta beraz (6.7.6) betetzen da. Orain m egoeran akzio pasiboa hartzen dela onartuko da eta $m-1$ egoeran aktibo akzioa, orduan hurrengoa betetzen da

$$\tilde{c}m + \lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0) + \theta m V_t^L(m-1,0) + \left(\lambda\frac{m}{L} + \mu + \theta(L-m)\right)V_t^L(m,0)$$
$$\geq \tilde{c}(m-1) + \min\Bigg(\lambda\left(1-\frac{m-1}{L}\right)V_t^L(m,0) + \left(\lambda\frac{m-1}{L} + \mu + \theta(L-m+1)\right)V_t^L(m-1,0)$$
$$\geq \tilde{c}(m-1) + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0),$$

non lehenengo inekuazioa goian frogatu den, eta bigarren inekuazioa m-n akzio optimoa aktiboa izatetik ondorioztatzen den. Orduan, (6.7.6). Inekuazioa betetzen da $m-1$ egoeran akzio aktiboa hartzen bada eta m egoeran akzio pasiboa. Azkenik, m egoeran aktibo akzioa eta $m-1$ egoeran akzio pasiboa deneko kasua falta da. Kasu honetan

$$\begin{aligned}
&\tilde{c}m + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0) \geq \tilde{c}(m-1) + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0)\\
&\geq \tilde{c}(m-1) + \min\left(\lambda\left(1 - \frac{m-1}{L}\right)V_t^L(m,0) + \left(\lambda\frac{m-1}{L} + \mu + \theta(L-m+1)\right)V_t^L(m-1,0)\right.\\
&+ \theta(m-1)V_t^L((m-2)^+,0),
\end{aligned}$$

non lehenengo inekuazioa goian frogatu den eta bigarren inekuazioa $m-1$-en akzio pasiboa optimo izatetik ondorioztatzen den. Orduan, (6.7.6). Inekuazioa betetzen da.

Orduan edozein t bada, $V_t(m,0)$ ez beherakorra da m-n. Eta $V_t^L(m,0) \to V^L(m,0)$ puntuz-puntu, orduan $V^L(m,0)$ ez-beherakorra da.

Orain $f_t^L(m,0) - f_t^L(m,1) \leq f_t^L(m',0) - f_t^L(m',1)$ frogatuko da non

$$\begin{aligned}
f_t^L(m,0) =&\lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0) + \theta m V_t^L(m-1,0) + \left(\lambda\frac{m}{L} + \mu + \theta(L-m)\right)V_t^L(m,0),\\
f_t^L(m,1) =&\tilde{c}m + c_s + \lambda V_t^L(1,1) + \theta L V_t^L(0,1) + \mu V_t^L(0,0).
\end{aligned}$$

$f_t^L(m,0) - f_t^L(m,1) \leq f_t^L(m',0) - f_t^L(m',1)$ **denaren forgapena edozein** $m' \geq m$**.** $f_t^L(m,a)$-ren espresioa ordezkatuz inekuazioa

$$\begin{aligned}
&\tilde{c}(m'-m) \leq \tilde{c}(m'-m) + \lambda\left(1-\frac{m'}{L}\right)V_t^L(m'+1,0)\\
&+\left(\lambda\frac{m'}{L} + \mu + \theta(L-m')\right)V_t^L(m',0) + \theta m' V_t^L(m'-1,0)\\
&-\left(\lambda\left(1-\frac{m}{L}\right)V_t^L(m+1,0) + \left(\lambda\frac{m}{L} + \mu + \theta(L-m)\right)V_t^L(m,0) + \theta m V_t^L(m-1,0)\right),
\end{aligned}$$

idatz daiteke. Definitu $m' = m + u$, non $u \geq 1$, eta orduan azken inekuazioa

$$\begin{aligned}
0 \leq&\lambda\left(1-\frac{m'}{L}\right)(V_t^L(m'+1,0) - V_t^L(m+1,0)) - \lambda\frac{u}{L}(V_t^L(m+1,0) - V_t^L(m,0))\\
&+\lambda\frac{m'}{L}(V_t^L(m',0) - V_t^L(m,0)) + \mu(V_t^L(m',0) - V_t^L(m,0)) + \theta m(V_t^L(m'-1,0) - V_t^L(m-1,0))\\
&+\theta u(V_t^L(m'-1,0) - V_t^L(m,0)),
\end{aligned}$$

idatz daiteke, zeina $V_t^L(m,0)$ ez-beherakorra izateagatik betetzen den.

6.7.2 6.2. Proposizioaren frogapena

Lehenik eta behin $\pi(m,1)$-ren espresioa garatuko da eta ondoren $\pi(m,0)$-ren espresioa.

Zerbitzaria lanpetuta dagoeneko oreka-egoerako banaketa

Lehenik eta behin $\pi(m,1)$ probabilitateari dagokion funtzio sortzailea definituko da edozein $m \in \mathbb{N}_0$ bada, hau da, $\Pi_1(z) = \sum_{m=0}^{\infty} z^m \pi(m,1)$, eta gogoratu (6.4.2). Ekuazioa edozein $m \in \mathbb{N}$ denean. Orduan, (6.4.2).

Ekuazioa m egoerarako z^m gaiaz biderkatuz, hau da

$$z^m(\lambda + m\theta + \mu)\pi(m,1) = z^m\lambda\pi(m-1,1) + z^m(m+1)\theta\pi(m+1,1), \quad \forall\, m \in \mathbb{Z}\backslash\{0\},$$

eta azken hau $m \in \{1,2,\ldots\}$ batuz, hurrengoa lor daiteke

$$(\lambda+\mu)\left(\Pi_1(z) - \pi(0,1)\right) + \theta z\frac{\mathrm{d}}{\mathrm{d}z}\Pi_1(z) = \lambda z\Pi_1(z) + \theta\left(\frac{\mathrm{d}}{\mathrm{d}z}\Pi_1(z) - \pi(1,1)\right).$$

Hainbat kalkuluren ostean azken hau

$$\frac{(\lambda(1-z)+\mu)\Pi_1(z)}{-\theta(1-z)} + \frac{\mathrm{d}\Pi_1(z)}{\mathrm{d}z} = \frac{(\lambda+\mu)\pi(0,1) - \theta\pi(1,1)}{-\theta(1-z)}, \tag{6.7.7}$$

idatz daiteke. Orain ekuazio diferentzial hau askatuko da. Hori egin ahal izateako, definitu $\Pi_1^n(z) = f_1(z)g_1(z)$ zeinetarako

$$\frac{\frac{\mathrm{d}f_1(z)}{\mathrm{d}z}}{f_1(z)} = -\frac{\lambda(1-z)+\mu}{-\theta(1-z)} \Rightarrow f_1(z) = \mathrm{e}^{\frac{\lambda z}{\theta}}(1-z)^{-\frac{\mu}{\theta}}. \tag{6.7.8}$$

$$\Pi_1(z) = f_1(z)g_1(z) = \frac{\mathrm{e}^{\lambda z/\theta}}{(1-z)^{\mu/\theta}}g_1(z),$$

espresioa (6.7.7). Ekuazioan ordezkatuz eta bi aldeak $-\theta(1-z)\frac{\mathrm{e}^{\lambda z/\theta}}{(1-z)^{\mu/\theta}}$ gaiaz zatiztuz

$$\frac{\mathrm{d}g_1(z)}{\mathrm{d}z} = \frac{(\lambda+\mu)\pi(0,1) - \theta\pi(1,1)}{-\theta(1-z)\mathrm{e}^{\lambda z/\theta}(1-z)^{-\mu/\theta}},$$

lor daiteke. Azken ekuazio hau integratuz, eta ohartuz, $f_1(0) = 1$ eta $\Pi_1(0) = \pi(0,1)$ direla, orduan $g_1(0) = \pi(0,1) \neq 0$, eta

$$\begin{aligned} g_1(z) &= \pi(0,1) - \int_0^z \frac{(\lambda+\mu)\pi(0,1) - \theta\pi(1,1)}{\theta\mathrm{e}^{\lambda x/\theta}(1-x)^{1-\mu/\theta}}\mathrm{d}x \\ \Longrightarrow g_1(z) &= \pi(0,1) - \frac{(\lambda+\mu)\pi(0,1) - \theta\pi(1,1)}{\theta}\int_0^z \frac{(1-x)^{\mu/\theta}}{\mathrm{e}^{\lambda x/\theta}(1-x)}\mathrm{d}x. \end{aligned} \tag{6.7.9}$$

Orain $\pi(m,1) = \frac{1}{m!}\frac{\mathrm{d}^m\Pi_1(z)}{\mathrm{d}z^m}|_{z=0}$-ren espresioa esplizitu bat garatzea da helburua edozein $m \geq 0$. (6.7.8) eta (6.7.9) ekuazioak kontutan hartuz

$$\Pi_1(z) = \frac{\pi(0,1)\mathrm{e}^{\frac{\lambda z}{\theta}}}{(1-z)^{\frac{\mu}{\theta}}} - \frac{((\lambda+\mu)\pi(0,1) - \theta\pi(1,1))}{\theta(1-z)^{\frac{\mu}{\theta}}}\mathrm{e}^{\frac{\lambda}{\theta}(z-1)} \cdot (-1)^{\frac{\mu}{\theta}-1}\left(\frac{\theta}{\lambda}\right)^{\frac{\mu}{\theta}}\int_{-\frac{\lambda}{\theta}}^{-\frac{\lambda}{\theta}(1-z)} y^{\frac{\mu}{\theta}-1}\mathrm{e}^{-y}\mathrm{d}y, \tag{6.7.10}$$

da, non $y = -\frac{\lambda}{\theta}(1-x)$ aldagai aldaketa erabili den integralean. Ohartu $\Pi_1(z)$-ren espresioan agertzen den integrala gamma funtzio ez-osoa dela [2, 6. Kapitulua]. Beraz, $\mu/\theta > 0$ denez,

$$\int_{-\frac{\lambda}{\theta}}^{-\frac{\lambda}{\theta}(1-z)} y^{\frac{\mu}{\theta}-1}\mathrm{e}^{-y}\mathrm{d}y = \left(-\frac{\lambda}{\theta}\right)^{\frac{\mu}{\theta}}\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i((1-z)^{i+\frac{\mu}{\theta}} - 1)}{i!(\frac{\mu}{\theta}+i)}. \tag{6.7.11}$$

$\pi(m,1)$ probabilitateak garatu aurretik edozein $m \in \mathbb{N}_0$ denean, ohartu $\Pi_1(z)$ ez dela ongi definitutako funtzio bat $z=1$ denean, eta beraz, $\lim_{z\to 1}\Pi_1(z)$ $0/0$ indeterminazio bat izatera bultzatu behar da. $z\to 1$ doan heinean (6.7.10). ekuazioko izendatzailea 0 izatea bultzatuz, hurrengo baldintza lortzen da

$$\pi(0,1)\mathrm{e}^{\lambda/\theta} - \left(\frac{\lambda+\mu}{\theta}\pi(0,1) - \pi(1,1)\right)\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)} = 0.$$

Azken ekuazio hau ebatziz $\pi(1,1)$-ren espresio esplizitua lortzen da $\pi(0,1)$-rekiko, hau da,

$$\pi(1,1) = a_1\pi(0,1), \quad a_1 = \frac{\lambda+\mu}{\theta} - \frac{\mathrm{e}^{\lambda/\theta}}{\sum_{i=0}^{\infty}\frac{(\lambda/\theta)^i}{i!(\mu/\theta+i)}}.$$

$\pi(1,1) = a_1\pi(0,1)$ ordezkatuz eta (6.7.11) (6.7.9). ekuazioan ordezkatuz, (6.7.8) eta (6.7.9) ekuazioetatik

$$f_1(z) = \mathrm{e}^{\frac{\lambda z}{\theta}}(1-z)^{-\frac{\mu}{\theta}},$$
$$g_1(z) = \pi(0,1)\left(1 + \frac{\sum_{j=0}^{\infty}\frac{(\frac{\lambda}{\theta})^j((1-z)^{j+\frac{\mu}{\theta}}-1)}{j!(\frac{\mu}{\theta}+j)}}{\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)}}\right),$$

lortzen da. Orain zerbitzaria lanean ari deneko oreka-egoerako banaketa kalkulatu daiteke, hau da, $\pi(m,1)$ edozein $m \geq 1$. Definitu $\ell_i(\mu/\theta) = \mu/\theta \cdot \ldots \cdot (\mu/\theta + i - 1)$ edozein $i \geq 1$ bada eta $\ell_0(\mu/\theta) = 1$ eta ohartu $\pi(m,1) = \frac{1}{m!}\frac{\mathrm{d}^m\Pi_1(z)}{\mathrm{d}z^m}|_{z=0} = \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}f_1^{(m-k)}g_1^{(k)}$ dela, non

$$f_1^{(k)} := \frac{\mathrm{d}^k f_1(z)}{\mathrm{d}z^k}|_{z=0} = \sum_{i=0}^{k}\binom{k}{i}\left(\frac{\lambda}{\theta}\right)^{k-i}\frac{\mathrm{e}^{\frac{\lambda z}{\theta}}\ell_i\left(\frac{\mu}{\theta}\right)}{(1-z)^{\frac{\mu}{\theta}+i}}|_{z=0} = \sum_{i=0}^{k}\binom{k}{i}\left(\frac{\lambda}{\theta}\right)^{k-i}\ell_i\left(\frac{\mu}{\theta}\right), \text{ for all } k \geq 0,$$

$$g_1^{(k)} := \frac{\mathrm{d}^k g_1(z)}{\mathrm{d}z^k}|_{z=0} = \pi(0,1)\frac{\sum_{j=0}^{\infty}\frac{\left(\frac{\lambda}{\theta}\right)^j\ell_k(-\frac{\mu}{\theta}-j)(1-z)^{j+\frac{\mu}{\theta}-k}}{j!(\frac{\mu}{\theta}+j)}}{\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)}}|_{z=0} = \pi(0,1)\frac{\sum_{j=0}^{\infty}\frac{\left(\frac{\lambda}{\theta}\right)^j\ell_k(-\frac{\mu}{\theta}-j)}{j!(\frac{\mu}{\theta}+j)}}{\sum_{i=0}^{\infty}\frac{(\frac{\lambda}{\theta})^i}{i!(\frac{\mu}{\theta}+i)}}, \text{for all } k \geq 1,$$

eta $g_1^{(0)} = \pi(0,1)$. Definitu $a_1(0) := 1, a_1(1) := a_1$ eta $a_1(m) := \frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}f_1^{(m-k)}g_1^{(k)}$ edozein $m \geq 2$. Orduan $\pi(m,1) = a_1(m)\pi(0,1)$ lortzen da non $a_1(m)$ 6.2. Proposizioan definitu den.

Oreka-egoerako banaketa zerbitzaria lanpetuta ez daogenean

Lehenik eta behin $\pi(m,0)$-ri dagokion funtzio sortzailea definituko da edozein $0 \leq m \leq n-1$ bada, hau da, $\Pi_0(z) = \sum_{m=0}^{\infty}z^m\pi(m,0) = \sum_{m=0}^{n-1}z^m\pi(m,0)$, non definizioz $\pi(m,0) = 0$ edozein $m \geq n$ bada, eta gogoratu (6.4.3). Ekuazioa edozein $1 \leq m \leq n-1$ denean. (6.4.3). Ekuazioa m egoeran z^m gaiaz biderkatuz, hau da

$$z^m(\lambda + m\theta)\pi(m,0) = z^m\lambda\pi(m-1,0) + z^m\mu\pi(m,1) + z^m(m+1)\theta\pi(m+1,0),$$

eta azken hau $1 \leq m \leq n-1$ batuz, hurrengoa lortzen da

$$\lambda(\Pi_0(z) - \pi(0,0)) + \theta z\frac{\mathrm{d}\Pi_0(z)}{\mathrm{d}z} = \theta\left(\frac{\mathrm{d}\Pi_0(z)}{\mathrm{d}z} - \pi(1,0)\right) + \lambda z(\Pi_0(z) - \pi(n-1,0)) + \mu\sum_{m=1}^{n-1}z^m\pi(m,1).$$

(6.4.3). Ekuazioa erabiliz $m = 0$ den kasuan, hau da, $\lambda\pi(0,0) - \theta\pi(1,0) = \mu\pi(0,1)$, eta hainbat kalkulu egin ostean, hurrengoa garatu daiteke

$$-\frac{\lambda}{\theta}\Pi_0(z) + \frac{\mathrm{d}\Pi_0(z)}{\mathrm{d}z} = \frac{\lambda z}{\theta(1-z)}\pi(n-1,0) - \frac{\mu}{\theta(1-z)}\sum_{m=0}^{n-1} z^m \pi(m,1). \tag{6.7.12}$$

Ohartu azken ekuazioan $\Pi_0(z)$ ongi definituta egon dadin $z = 1$-en, zeina $\sum_{m=0}^{n-1}\pi(m,0) < 1$ den, hurrengo baldintzak bete behar du

$$\lim_{z\to 1} \lambda z\pi(n-1,0) - \mu\sum_{m=0}^{n-1}\pi(m,1) = 0.$$

Orduan problemaren beste baldintza bat lortzen da, hau da,

$$\pi(n-1,0) = \frac{\mu}{\lambda}\sum_{m=0}^{n-1}\pi(m,1) = \frac{\mu}{\lambda}\pi(0,1)\sum_{m=0}^{n-1}a_1(m),$$

non $a_1(m)$ 6.2. Proposizioan definitu den. Azken honek $\pi(n-1,0) = a_0^n(n-1)\pi(0,1)$ betetzen du. $\pi(m,0)$-ren espresioa garatu ahal izateko edozein $n-2 \geq m \geq 1$, hurrengo oreka ekuazioa onartuko da edozein $n-2 \geq m \geq 1$bada, zeina (6.4.3). Ekuazioen baliokidea den:

$$\lambda\pi(m,0) = \theta(m+1)\pi(m+1,0) + \lambda\pi(n-1,0) - \mu\sum_{j=m+1}^{n-1}\pi(j,1).$$

Azken ekuazio hau 6.4.1. Sekzioan erabilitako antzeko argumentuak erabiliz ebatzi daiteke, eskuinaldeko lehenengo bi gaiak $\mu = \infty$-ren oreka ekuazioei baitagokie. Honetaz ohartuz, errekurtsioa erraz ebatzi daiteke $a_0^n(m)$-ren espresioa lortzeko edozein $n-2 \geq m \geq 1$ bada, ikusi 6.2. Proposizioa. Hori egin ahal izateko (6.4.3). Ekuazioa gogoratu, eta ohartu hurrengo ekuazioaren baliokidea dela edozein $n-2 \geq m \geq 1$ bada:

$$\lambda\pi(m,0) = \theta(m+1)\pi(m+1,0) + \lambda\pi(n-1,0) - \mu\sum_{j=m+1}^{n-1}\pi(j,1).$$

Orduan,

$$
\begin{aligned}
\pi(m,0) =&\frac{\theta(m+1)}{\lambda}\pi(m+1,0)+\pi(n-1,0)-\frac{\mu}{\lambda}\sum_{j=m+1}^{n-1}\pi(j,1),\\
=&\frac{\theta(m+1)}{\lambda}\left(\frac{\theta(m+2)}{\lambda}\pi(m+2,0)+\pi(n-1,0)-\frac{\mu}{\lambda}\sum_{j=m+2}^{n-1}\pi(j,1)\right)+\pi(n-1,0)-\frac{\mu}{\lambda}\sum_{j=m+1}^{n-1}\pi(j,1)\\
=&\frac{\theta^2}{\lambda^2}(m+1)(m+2)\pi(m+2,0)+\left(1+\frac{\theta}{\lambda}(m+1)\right)\pi(n-1,0)-\frac{\mu}{\lambda}\pi(m+1,1)\\
&-\frac{\mu}{\lambda}\sum_{j=m+2}^{n-1}\left(1+\frac{\theta(m+1)}{\lambda}\right)\pi(j,1)\\
=&\dots\\
=&\pi(n-1,0)\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}-\frac{\mu}{\lambda}\sum_{j=m+1}^{n-1}\pi(r,1)\sum_{i=0}^{r-m-1}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}.
\end{aligned}
$$

Azken espresio hau edozein $m \geq 1$-rako da baliogarria.

Azkenik, $a_0^n(0)$-ren espresioa $\pi(0,0)=\frac{\mu}{\lambda}\pi(0,1)+\frac{\theta}{\lambda}a_0^n(1)\pi(0,1)$ askatuz lor daiteke. Orduan, $\pi(m,0)=a_0^n(m)\pi(0,1)$ lortzen da.

6.7.3 6.3. Proposizioaren frogapena

Helburua edozein $n' \neq n$ denean

$$\tilde{c}\mathbb{E}(N^n)+c_sP_b^n \leq \tilde{c}\mathbb{E}(N^{n'})+c_sP_b^{n'},$$

dela frogatzea da. Hemen $n' < n$ kasurako frogapena aurkeztuko da, beste kasua antzeko modura egin daiteke. Hipotesiz $n \geq 1$ denean $\alpha(n-1) \leq \alpha(n)$, orduan

$$
\begin{aligned}
&\frac{\mathbb{E}(N^{n-1})-\mathbb{E}(N^{n-2})}{P_b^{n-2}-P_b^{n-1}} \leq \frac{\mathbb{E}(N^n)-\mathbb{E}(N^{n-1})}{P_b^{n-1}-P_b^n}\\
&\Longrightarrow (\mathbb{E}(N^{n-1})-\mathbb{E}(N^{n-2}))(P_b^{n-1}-P_b^n) \leq (\mathbb{E}(N^n)-\mathbb{E}(N^{n-1}))(P_b^{n-2}-P_b^{n-1}).
\end{aligned}
$$

Azken inekuazioan $\mathbb{E}(N^n)(P_b^{n-1}-P_b^n)$ batu eta kenduko da ezkerraldean, hau da,

$$(\mathbb{E}(N^{n-1})-\mathbb{E}(N^n)+\mathbb{E}(N^n)-\mathbb{E}(N^{n-2}))(P_b^{n-1}-P_b^n) \leq (\mathbb{E}(N^n)-\mathbb{E}(N^{n-1}))(P_b^{n-2}-P_b^{n-1}),$$

hainbat kalkulu egin ostean azken inekuazio hau

$$\frac{\mathbb{E}(N^n)-\mathbb{E}(N^{n-2})}{P_b^{n-2}-P_b^n} \leq \frac{\mathbb{E}(N^n)-\mathbb{E}(N^{n-1})}{P_b^{n-1}-P_b^n} \leq c_s,$$

idatz daiteke. Antzera, frogatu daiteke

$$\alpha(n-1) \leq \frac{\mathbb{E}(N^n)-\mathbb{E}(N^{n-2})}{P_b^{n-2}-P_b^n}.$$

Orain hurrengo indukzio hipotesia egingo da a finko batentzat

$$\alpha(n-a+1) \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a})}{P_b^{n-a} - P_b^n} \leq c_s.$$

Enuntziatuko hipotesitik $\alpha(n-a) \leq \alpha(n-a+1)$ lortzen da, beraz azken ekuaziotik

$$\begin{aligned} &\alpha(n-a) \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a})}{P_b^{n-a} - P_b^n} \\ &\Longrightarrow (\mathbb{E}(N^{n-a}) - \mathbb{E}(N^{n-a-1}))(P_b^{n-a} - P_b^n) \leq (\mathbb{E}(N^n) - \mathbb{E}(N^{n-a}))(P_b^{n-a-1} - P_b^{n-a}), \end{aligned}$$

lortzen da, $\mathbb{E}(N^n)(P_b^{n-a} - P_b^n)$ gaia ezkerraldean batu eta kenduz, eta hainbat kalkulu egin ostean,

$$\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a-1})}{P_b^{n-a-1} - P_b^n} \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a})}{P_b^{n-a} - P_b^n} \leq c_s, \tag{6.7.13}$$

lortzen da. Azken honetatik ohartu

$$(\mathbb{E}(N^n) - \mathbb{E}(N^{n-a-1}))(P_b^{n-a} - P_b^n + P_b^{n-a-1} - P_b^{n-a-1}) \leq (\mathbb{E}(N^n) - \mathbb{E}(N^{n-a}))(P_b^{n-a-1} - P_b^n),$$

zeina hainbat kalkulu egin ostean

$$\alpha(n-a) \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a-1})}{P_b^{n-a-1} - P_b^n},$$

idatz daitekeen. Azken hau (6.7.13). ekuazioarekin batera

$$\alpha(n-a) \leq \frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-a-1})}{P_b^{n-a-1} - P_b^n} \leq c_s,$$

da, zeinak indukzio argumentua amaitzen duen. $0 \leq a \leq n-1$ den kasuan definitu $n' = n-1-a$. Edozein $n' < n$-rako hurrengoa frogatu da

$$\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n'})}{P_b^{n'} - P_b^n} \leq c_s \Longrightarrow \tilde{c}\mathbb{E}(N^n) + c_s P_b^n \leq \tilde{c}\mathbb{E}(N^{n'}) + c_s P_b^{n'}.$$

Zeinak frogapena amaitzen duen.

6.7.4 6.1. Lemaren frogapena

Lehenik eta behin P_b^n ez-beherakorra dela frogatuko da. Orduan nahikoa da $\pi_{n-1}^n \leq \pi_{n-2}^{n-1}$ frogatzea edozein $n \geq 2$ denean. Azken inekuazio hau

$$\sum_{m=0}^{n-2} \sum_{i=0}^{n-2-m} \left(\frac{\theta}{\lambda}\right)^i \frac{(m+i)!}{m!} \leq \sum_{m=0}^{n-1} \sum_{i=0}^{n-1-m} \left(\frac{\theta}{\lambda}\right)^i \frac{(m+i)!}{m!} \Leftrightarrow 0 \leq \sum_{m=0}^{n-1} \left(\frac{\theta}{\lambda}\right)^{n-1-m} \frac{(n-1)!}{m!},$$

idatz daiteke. Eskuinaldea positiboa da edozein $n \geq 2$ denean eta beraz $\pi_{n-1}^n \leq \pi_{n-2}^{n-1}$. π_{n-1}^n n-n ez-beherakorra dela frogatu ostean, P_b^n ganbila dela frogatu nahi da. P_b^n ganbila izatea $\pi_{n-1}^n - \pi_n^{n+1} \leq$

$\pi_{n-2}^{n-1} - \pi_{n-1}^{n}$ inekuazioak inplikatzen du, zeina hainbat kalkuloren ostean

$$\begin{aligned}
&\frac{\sum_{m=0}^{n+1}\sum_{i=0}^{n+1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!} - \sum_{m=0}^{n}\sum_{i=0}^{n-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}}{\sum_{m=0}^{n+1}\sum_{i=0}^{n+1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}}\\
&\leq \frac{\sum_{m=0}^{n}\sum_{i=0}^{n-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!} - \sum_{m=0}^{n-1}\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}}{\sum_{m=0}^{n-1}\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}},
\end{aligned} \tag{6.7.14}$$

idatz daitekeen. Hainbat kalkuloren ostean hurrengoa ondorioztatzen da:

$$\begin{aligned}
&\left(\sum_{m=0}^{n+1}\left(\frac{\theta}{\lambda}\right)^{n+1-m}\frac{(n+1)!}{m!}\right)\sum_{m=0}^{n-1}\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\\
&\leq \left(\sum_{m=0}^{n}\left(\frac{\theta}{\lambda}\right)^{n-m}\frac{n!}{m!}\right)\sum_{m=0}^{n+1}\sum_{i=0}^{n+1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}.
\end{aligned}$$

Azken hau ezin izan da formalki frogatu, halere, azken inekuazioa numerikoki kalkulatu da n-ren balio gorakorrentzat eta λ eta θ-ren zoriozko balioak hartuz, zeineratako ikusi den inekuazioa betetzen dela. Ondoren analitikoki kalkulatu da inekuazioa betetzen dela n infiniturantz doanean. Beraz, P_b^n ganbila da n-n.

Froga dezagun orain 6.3. Proposizioan definitu den $\alpha(n)$ funtzioa ez-beherakorra dela,

$$\begin{aligned}
\alpha(n) &:= \tilde{c}\frac{\mathbb{E}(N^n) - \mathbb{E}(N^{n-1})}{P_b^{n-1} - P_b^n} = \frac{\sum_{m=0}^{n-1} m\pi_m^n - \sum_{m=0}^{n-2} m\pi_m^{n-1}}{P_b^{n-1} - P_b^n}\\
&= \frac{\sum_{m=0}^{n-1} m\left(\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\right)\pi_{n-1}^n}{P_b^{n-1} - P_b^n} - \frac{\sum_{m=0}^{n-2} m\left(\sum_{i=0}^{n-2-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\right)\pi_{n-2}^{n-1}}{P_b^{n-1} - P_b^n}.
\end{aligned}$$

Azken hau hainbat kalkuloren ostean

$$\alpha(n) = \frac{\sum_{m=0}^{n-1} m\left(\frac{\theta}{\lambda}\right)^{n-1-m}\frac{(n-1)!}{m!}\pi_{n-1}^n}{P_b^{n-1} - P_b^n} + \sum_{m=0}^{n-2} m\sum_{i=0}^{n-2-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\left(\frac{\pi_{n-1}^n - \pi_{n-2}^{n-1}}{P_b^{n-1} - P_b^n}\right),$$

idatz daiteke. Orain $\alpha(n) \leq \alpha(n+1)$ dela frogatu nahi da edozein n denean, honetarako

$$\begin{aligned}
&\frac{\sum_{m=0}^{n-1} m\left(\frac{\theta}{\lambda}\right)^{n-1-m}\frac{(n-1)!}{m!}\pi_{n-1}^n}{P_b^{n-1} - P_b^n} + \sum_{m=0}^{n-2} m\sum_{i=0}^{n-2-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\left(\frac{\pi_{n-1}^n - \pi_{n-2}^{n-1}}{P_b^{n-1} - P_b^n}\right)\\
&\leq \frac{\sum_{m=0}^{n} m\left(\frac{\theta}{\lambda}\right)^{n-m}\frac{n!}{m!}\pi_n^{n+1}}{P_b^n - P_b^{n+1}} + \sum_{m=0}^{n-1} m\sum_{i=0}^{n-1-m}\left(\frac{\theta}{\lambda}\right)^i\frac{(m+i)!}{m!}\left(\frac{\pi_n^{n+1} - \pi_{n-1}^n}{P_b^n - P_b^{n+1}}\right),
\end{aligned}$$

bete behar da. Bigarren gaia inekuazioren ezkerraldean eskuinaldeko bigarren gaia baino txikiagoa edo berdina da, hori P_b^n eta π_{n-1}^n funtzio ganbila eta ez-beherakorra izateak inplikatzen dute. Beraz,

$$\sum_{m=0}^{n-1} m\left(\frac{\theta}{\lambda}\right)^{n-1-m}\frac{(n-1)!}{m!}\pi_{n-1}^n \leq \sum_{m=0}^{n} m\left(\frac{\theta}{\lambda}\right)^{n-m}\frac{n!}{m!}\pi_n^{n+1},$$

frogatzearekin nahikoa da. Azken hau ezin izan da formalki frogatu, halere, azken ineka azken inekuazioa numerikoki kalkulatu da n-ren balio gorakorrentzat eta λ eta θ-ren zoriozko balioak hartuz, zeineratako ikusi den inekuazioa betetzen dela. Ondoren analitikoki kalkulatu da inekuazioa betetzen dela n infiniturantz doanean.

6.7.5 6.2. Lemaren frogapena

P_b^n ez-beherakorra dela ikusteko nahikoa da $\pi^n(0,1) \geq \pi^{n+1}(0,1)$ ikustea. Ohartu $\pi^n(0,1) \geq \pi^{n+1}(0,1)$, $\sum_{m=0}^{n-1} a_0^n(m) \leq \sum_{m=0}^{n} a_0^{n+1}(m)$ inekuazioak inplikatzen duela. Ez hori bakarrik, hainbat kalkuloren ostean batek

$$\sum_{m=0}^{n} a_0^{n+1}(m) = \sum_{m=0}^{n} a_0^n(m) + \frac{\mu}{\lambda}\sum_{r=0}^{n} a_1(r)\left(\frac{\theta}{\lambda}\right)^{n-m}\frac{n!}{m!},$$

lor dezake, orduan

$$0 \leq \sum_{m=0}^{n} a_0^{n+1}(m) - \sum_{m=0}^{n-1} a_0^n(m) \Longleftrightarrow 0 \leq a_0^n(n) + \frac{\mu}{\lambda}\sum_{r=0}^{n} a_1(r)\left(\frac{\theta}{\lambda}\right)^{n-m}\frac{n!}{m!}.$$

Azken inekuazio hau $a_0^n(m) > 0$ edozein m-tarako izateak inplikatzen du.

A. Eranskina

Eranskina

Eranskin honetan kontrol determinista eta kontrol estokastikoari dagokien emaitza esanguratsuenak aurkeztuko dira. A.1. Sekzioan politika egonkor optimo baten existentzia bermatu dadin baldintzak eskeiniko dira helburua batez besteko kostua minimizatzea denean. A.2. Sekzioan Prontyagin-en Minimoaren Printzipioa azalduko da, hau da, optimoa izan dadin beharrezko baldintzak kontrol optimoko problema determinista batean. A.3. Sekzioan soluzio bat optimoa izan dadin baldintza nahikoak aurkeztuko dira.

A.1 Politika optimo egonkorren existentziareko baldintzak: Bellman ekuazioa

Gogoratu (1.3.4)-ko helburu funtzioa, eta definitu

$$\mathcal{C}^{\phi} = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T C(N^{\phi}(t), S^{\phi}(N^{\phi}(t)))\mathrm{d}t\right).$$

Politika optimo egonkor bat existitu dadin baldintzak hurrengoak dira.

A.1 Teorema. *Existitzen badira $V(\cdot)$ eta $g \in \mathbb{R}$ zeinentzat*

$$g + V(m) = \min_{a\in\mathcal{A}}\{C(m,a) + \sum_{\tilde{m}=0}^{\infty} p^a(\tilde{m}, m)V(\tilde{m})\},$$

orduan existitzen da ϕ^*, politika estazionario bat zeinentzat

$$g = \mathcal{C}^{\phi^*} = \min_{\phi} \mathcal{C}^{\phi}.$$

Informazio gehigarria [53, 76, 79] liburuetan aurki daiteke.

A.2 Beharrezko baldintzak: Pontryagin-en Printzipio minimioa

Tesi honen sarrera aipatu den moduan, beharrezko baldintzak soluzio bat optimoa izan dadin lagungarriak dira optimoak izan daitezkeen aukeragaiak aurkitzeko. Ostean, baldintza nahikoak erabil daitezke hautagai horiek optimoak diren ala ez ziurtatzeko [83], edo elkarren arteak erkatu daitezke euretatik optimoa zein den kalkulatzeko.

Gogoratu Hamiltondarra eta Lagrangearraren definizioak, 1.3.4. Sekzioan definitu bezala, hau da

$$\mathcal{H}(m(t), s(t), \gamma(t)) := C(m(t), s(t)) + \gamma^{\mathsf{T}}(t) f(m(t), s(t)), \text{ eta,}$$

$$\mathcal{L}(m(t), s(t), \gamma(t), \nu(t), \omega(t)) := \mathcal{H}(m(t), s(t), \gamma(t)) + \nu^{\mathsf{T}}(t) h_1(s(t)) + \omega^{\mathsf{T}}(t) h_2(m(t)).$$

Izan bedi $\mathcal{H}(t) := \mathcal{H}(m(t), s(t), \gamma(t))$ eta $\mathcal{L}(t) := \mathcal{L}(m(t), s(t), \gamma(t), \nu(t), \omega(t))$. Orduan hurrengo teoremak $(m(t), s(t))$ bikotea optimoa izan dadin beharrezko baldintzak eskeintzen ditu. Ikusi [55] ikerketa artikulua Minimoaren Printzipioaren enuntziatu orokorragoak ikusteko.

A.2 Teorema. *Izan bedi $s^*(\cdot)$ kontrol optimoa, zeina zatika jarraia den, eta izan bedi T amaierako denbora optimoa (zeina optimizazioaren emaitza den) eta $m^*(\cdot)$ berari dagokion ibilbide optimoa $[0, T]$ tartean. Orduan, existitzen da $\gamma^*(t) = (\gamma_1^*(t), \ldots, \gamma_K^*(t))$ funtzio adjuntoa zatika jarraia dena eta zeinaren deribatua zatika jarraia den, zatika jarraiak diren biderkatzaileak $\nu(t)$ eta $\omega(t)$ edozein $t \in [0, T]$ tartean, zeinak*

1. $$\dot{\gamma}^* = -\frac{\partial \mathcal{L}(t)}{\partial m^*}, \tag{A.2.1}$$

   betetzen duen edozein $m^*(t)$-ren puntu jarraietan, non $\mathcal{L}$ sistemaren Lagrangearra den.

2. $$s^*(t) = \arg\min_{s(t) \in \mathcal{S}} \mathcal{H}(m^*(t), s^*(t), \gamma^*(t)), \tag{A.2.2}$$

   non $\mathcal{U}$, onargarriak diren kontrolen multzoa den. Orduan, azken balditza

   $$\frac{\partial \mathcal{L}^*(t)}{\partial s^*} = 0, \tag{A.2.3}$$

   idatz daiteke

3. $$\dot{m}^*(t) = f(m^*(t), s^*(t)), \tag{A.2.4}$$

   $s^*(t)$-ren puntu jarrai guztietan, eta $m^*(0) = m_0$.

4. *Ez hori bakarrik,*

   $$\nu(t) h_1(s^*(t)) = 0, \nu(t) \geq 0 \text{ and } \omega(t) h_2(m^*(t)) = 0, \omega(t) \geq 0 \text{ edozein } t \in [0, T]. \tag{A.2.5}$$

Gainera, existitzen da K_0 konstantea zeinak

$$\mathcal{H}(m^*(t), s^*(t), \gamma^*(t)) = K_0, \text{ edozein } t \in [0, T]. \tag{A.2.6}$$

betetzen duen.

K_0 konstantea 0 da amaierako denbora akea denean.

A.3 Baldintza nahikoak: Hamilton-Jacobi-Bellman Ekuazioa

Sekzio honetan kontrol optimo problema baten soluzio optimoa izan dadin baldintza nahikoak aurkeztuko dira.

A.3 Teorema. *Demagun existitzen dela $V(t, m)$ funtzio jarraiki diferentziagarria t-n eta m-n zeinetarako*

$$0 = \min_{s \in \mathcal{S}} [C(m, s) + \nabla_t V(t, m) + \nabla_m V(t, m)^{\mathsf{T}} f(m, s)], \text{ edozein } t, m. \tag{A.3.1}$$

$s^*(t)$ kontrol traiektoriak (1.3.11) ekuazioaren eskuinaldea minimizatzen badu edozein $t \in [0, T]$ denean, $s^*(t)$ zatika jarraia bada t-n eta $\frac{\mathrm{d}m(t)}{\mathrm{d}t} = f(m^*(t), s^*(t))$ emaitza bakarra badu (t, m) bikoterentzat, non $m^*(t)$ $s^*(t)$-ri dagokion traiektoria den. Orduan, $V(t, m)$ kostu optimoa da edozein t, m bada eta kontrol traiektoria $s^*(x)$ edozein $t \in [0, T]$-rako optimoa da.

Autorearen artikuluak

[SR1] M. Larrañaga, O.J. Boxma, R. Núñez-Queija and M.S. Squillante. “Efficient content delivery in the presence of impatient jobs". In: *Proceedings of 27th International Teletraffic Congress* (2015).

[SR2] M. Larrañaga, U. Ayesta and I.M. Verloop. “Asymptotically optimal index policies for an abandonment queue with convex holding cost". In: *Queueing Systems* (2015), pp.1–73.

[SR3] M. Larrañaga, U. Ayesta and I.M. Verloop. “Dynamic control of birth-and-death restless bandits: application to resource-allocation problems". *Submitted to IEEE/ACM Transactions on Networking* (2015).

[SR4] M. Larrañaga, U. Ayesta and I.M. Verloop. “Dynamic fluid-based scheduling in a multi-class abandonment queue". In: *Performance Evaluation 70.10, pp. 841–858, Special Issue of IFIP Performance (2013)*.

[SR5] M. Larrañaga, U. Ayesta and I.M. Verloop. “Index Policies for a multi-class queue with convex holding cost an abandonment". In: *Proceedings of ACM SIGMETRICS* (2014), pp.125–137.

[SR6] M. Larrañaga, U. Ayesta and I.M. Verloop. “Stochastic and fluid index policies for resource allocation problems". In: *Proceedings of IEEE Infocom* (2015).

Bibliografia

[1] S. Aalto, P. Lassila, and P. Osti. “Whittle Index Approach to Size-aware Scheduling with Time-varying Channels”. In: *Proceedings of ACM Sigmetrics* (2015), pp. 57–69 (cit. on p. 9).

[2] M. Abramowitz and A.I. Stegun. *Handbook of Mathematical Functions, with Formulas, Graphs and Mathematical Tables*. Dover Publications, Inc., New York, 1965 (cit. on p. 147).

[3] D.F. Anderson. *Introduction to Stochastic Processes with Applications in the Biosciences*. University of Wisconsin at Madison, 2013 (cit. on pp. 35, 134).

[4] P.S. Ansell, K.D. Glazebrook, I. Mitrani, and J. Niñ-Mora. “A semidefinite programming approach to the optimal control of a single server queueing system with imposed second moment constraints”. In: *Journal of the Operational Research Society* 50 (1999), pp. 765–773 (cit. on p. 6).

[5] P.S. Ansell, K.D. Glazebrook, J. Niñ-Mora, and M. O'Keeffe. “Whittle's index policy for a multi-class queueing system with convex holding costs”. English. In: *Mathematical Methods of Operations Research* 57.1 (2003), pp. 21–39 (cit. on pp. 4, 6, 10, 20, 24, 27).

[6] N.T. Argon, S. Ziya, and R. Righter. “Scheduling Impatient Jobs in a Clearing System with Insights on Patient Triage in Mass-casualty Incidents”. In: *Probability in the Engineering and Informational Sciences* 22.3 (2010), pp. 301–332 (cit. on p. 7).

[7] B. Ata and M.H. Tongarlak. “On scheduling a multiclass queue with abandonments under general delay costs”. In: *Queueing Systems* 74.1 (2013), pp. 65–104 (cit. on p. 7).

[8] R. Atar, C. Giat, and N. Shimkin. “The $c\mu/\theta$ rule for many-server queues with abandonment”. In: *Operation Research* 58.5 (2010), pp. 1427–1439 (cit. on pp. 7, 22, 39–41, 76, 106–108, 112, 113, 116).

[9] R. Atar, C. Giat, and N. Shimkin. “On the asymptotic optimality of the $c\mu/\theta$ rule under ergodic cost”. In: *Queueing Systems* 67 (2011), pp. 127–144 (cit. on pp. 7, 106–108, 112, 113, 116).

[10] R. Atar, A. Mandelbaum, and M.I. Reiman. “Scheduling a multi-class queue with many exponential servers: Asymptotic optimality in heavy-traffic”. In: *The Annals of Applied Probability* 14 (2004), pp. 1084–1134 (cit. on p. 106).

[11] K. Avrachenkov and V. Borkar. “Whittle Index Policy for Crawling Ephemeral Content”. In: *arXiv preprint arXiv: 1503.08558* (2015) (cit. on p. 9).

[12] F. Avram. “Optimal Control of Fluid Limits of Queueing Networks and Stochasticity Corrections”. In: *Mathematics of Stochastic Manufacturing Systems: AMS-SIAM Summer Seminar in Applied Mathematics* 33 (1996), pp. 1–36 (cit. on p. 10).

[13] F. Avram, D. Bertsimas, and M. Ricard. “An Optimal Control Approach to Optimization of Multiclass Queueing Networks”. In: *Proceedings of Workshop on Queueing Networks of the Mathematical Inst.* (1994) (cit. on p. 10).

[14] F. Avram, D. Bertsimas, and M. Richard. “Optimization of multiclass queuing networks: a linear control approach”. In: (1994), pp. 199–234 (cit. on pp. 10, 71).

[15] U. Ayesta, P. Jacko, and V. Novak. “A nearly-optimal index rule for scheduling of users with abandonment”. In: *Proceedings of IEEE Infocom* (2011) (cit. on pp. 7, 39–41, 107, 108, 116).

[16] F. Baccelli, P. Boyer, and G. Hebuterne. “Single-server queues with impatient customers”. In: *Advances in Applied Prabability* 16 (1984), pp. 887–905 (cit. on p. 7).

[17] N. Bäuerle. “Asymptotic optimality of tracking policies in stochastic networks”. In: *The Annals of Applied Probability* 10 (2000), pp. 1065–1083 (cit. on pp. 10, 113).

[18] N. Bäuerle and U. Rieder. “Optimal control of single-server fluid networks”. In: *Queueing Systems* 35 (2000), pp. 185–200 (cit. on p. 10).

[19] R. Bellman. *Dynamic Programming*. 1st ed. Princeton, NJ, USA: Princeton University Press, 1957 (cit. on p. 12).

[20] D.P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, 19982 (cit. on p. 9).

[21] D.P. Bertsekas. *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999 (cit. on p. 9).

[22] D.P. Bertsekas. *Dynamic programming and optimal control*. Athena Scientific, 2005 (cit. on pp. 8, 12, 13, 15, 90).

[23] D. Bertsimas and J. Niñ-Mora. “Restless Bandits, Linear Programming Relaxations, and a Primal-Dual Index Heuristic”. In: *Operations Research* 48 (2000), pp. 80–90 (cit. on p. 4).

[24] S. Bhulai, H. Blok, and F.M. Spieksma. “K competing queues with customer abandonment: optimality of a generalized $c\mu$-rule by the smoothed rate truncation method”. In: *Work in progress* (2014) (cit. on pp. 7, 36).

[25] S. Bhulai, A.C. Brooms, and F.M. Spieksma. “On structural properties of the value function for an unbounded jump Markov process with an application to a processor sharing retrial queue”. In: *Queueing Systems* 76 (2014), pp. 425–446 (cit. on pp. 12, 37, 52, 132, 135, 142, 143).

[26] C.F. Bispo. “The single-server scheduling problem with convex costs”. In: *Queueing Systems* 73 (2013), pp. 261–294 (cit. on pp. 20, 80).

[27] N.K. Boots and H. Tijms. “A multiserver queueing systems with impatient customers”. In: *Management Science* 45.3 (1999), pp. 444–448 (cit. on p. 7).

[28] S.C. Borst. “User-level performance of channel-aware scheduling algorithms in wireless data networks”. In: *IEEE/ACM Transactions on Networking* 13.3 (2005), pp. 636–647 (cit. on pp. 22, 81).

[29] O.J. Boxma and P.R. de Waal. “Multiserver queues with impatient customers”. In: *Proceedings of 14th International Teletraffic Congress* (1994), pp. 743–756 (cit. on p. 7).

[30] A. Brandt and M. Brandt. “On the two-class M/M/1 system under preemptive resume and impatience of the prioritized customers”. In: *Queueing Systems* 47 (2004), pp. 147–168 (cit. on p. 7).

[31] P.H. Brill and M.J.M. Posner. “Level crossings in point processes applied to queues: single-server case”. In: *Operations Research* 25.4 (1977), pp. 662–674 (cit. on p. 7).

[32] O. Brun, H. Ben Cheikh, and B. Prabhu. “Optimal speed scaling for multiclass queues”. In: *HAL report: hal-01094506* (2014) (cit. on p. 10).

[33] C. Buyukkoc, P. Varaya, and J. Walrand. “The $c\mu$ rule revisited”. In: *Advances in Applied Probability* 17 (1985), pp. 237–238 (cit. on pp. 20, 43, 106, 116).

[34] B.C. Chachuat. *Nonlinear and Dynamic Optimization: from theory to practice.* IC-32: Winter Semester, 2006-2007 (cit. on pp. 13, 120).

[35] J.W. Cohen. *The single server queue*. North-Holland, 1982 (cit. on p. 2).

[36] J.G. Dai. “On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models”. In: *Annals of Applied Probability* 5 (1995), pp. 49–77 (cit. on p. 10).

[37] J.G. Dai and S. He. “Many-server queues with customer abandonment: A survey of diffusion and fluid approximations”. In: *Journal of Systems Science and Systems Engineering* 21 (1 2012), pp. 1–36 (cit. on p. 7).

[38] J.G. Dai and Dacheng Yao. “Brownian inventory models with convex holding cost, Part1: average-optimal controls”. In: *Stochastic Systems* 3.2 (2013), pp. 442–499 (cit. on p. 6).

[39] J.G. Dai and Dacheng Yao. “Brownian inventory models with convex holding cost, Part2: discount-optimal controls”. In: *Stochastic Systems* 3.2 (2013), pp. 500–573 (cit. on p. 6).

[40] D.G. Down, G. Koole, and M.E. Lewis. “Dynamic control of a single server system with abandonments”. In: *Queueing Systems* 67 (2011), pp. 63–90 (cit. on pp. 7, 36, 107).

[41] B. Doytchinov, J. Lehoczky, and S. Shreve. “Real-time queues in heavy traffic with earliest-deadline-first queue discipline”. In: *The Annals of Applied Probability* 11.2 (2001), pp. 332–378 (cit. on p. 132).

[42] A.K. Erlang. “The Theory of Probabilities and Telephone Conversations”. In: *Nyt Tidsskrift for Matematik B* 20 (1909) (cit. on p. 2).

[43] A. Gajrat and A. Hordijk. “Fluid approximation of a controlled multiclass tandem network”. In: *Queueing Systems* 35 (2000), pp. 349–380 (cit. on pp. 10, 113).

[44] O. Garnett, A. Mandelbaum, and M. Reiman. “Designing a call center with impatient customers”. In: *Manufacturing and Service Operations Management* 4 (2002), pp. 208–227 (cit. on p. 106).

[45] E. Gelenbe and I. Mitrani. *Analysis and Synthesis of Computer Systems*. London: Academic Press, 1980 (cit. on pp. 20, 43).

[46] J.M. George and J.M. Harrison. “Dynamic control of a queue with adjustable service rate”. In: *Operations Research* 49.5 (2001), pp. 720–731 (cit. on p. 20).

[47] J.C. Gittins, K. Glazebrook, and R. Weber. *Multi-armed Bandit Allocation Indices*. Wiley, 2011 (cit. on pp. 3, 4, 9, 24, 27, 28, 43, 44).

[48] K.D. Glazebrook, C. Kirkbride, and J. Ouenniche. “Index policies for the admission control and routing of impatient customers to heterogeneous servcie stations”. In: *Operations Research* 57 (2009), pp. 975–989 (cit. on pp. 22, 24, 27, 28).

[49] K.D. Glazebrook, J. Niñ-Mora, and P.S. Ansell. “Index policies for a class of discounted restless bandits”. In: *Advances in Applied Probability* (2002), pp. 754–774 (cit. on p. 4).

[50] K.D. Glazebrook, P.S. Ansell, R.T. Dunn, and R.R. Lumley. “On the optimal allocation of service to impatient tasks”. In: *Journal of Applied Probability* 41.1 (2004), pp. 51–72 (cit. on p. 7).

[51] K.D. Glazebrook, D.J. Hodge, C. Kirkbride, and R.J. Minty. “Stochastic scheduling: A short history of index policies and new approaches to index generation for dynamic resource allocation”. In: *Journal of Scheduling* (2013), pp. 1–19 (cit. on p. 10).

[52] C. Gromoll, P. Robert, and B. Zwart. “Fluid Limits for Processor Sharing Queues with Impatience”. In: *Mathematics of Operations Research* 33.2 (2008), pp. 375–402 (cit. on p. 7).

[53] X. Guo and O. Hernandez-Lerma. *Continuous-Time Markov Decision Processes*. Vol. 62. Springer-Verlag, 2013 (cit. on pp. 4, 11, 155).

[54] J.M. Harrison and A. Zeevi. “Dynamic scheduling of a multiclass queue in the Halfin-Whitt heavy traffic regime”. In: *Operations Research* 52.2 (2004), pp. 243–257 (cit. on p. 106).

[55] R. F. Hartl, S. P. Sethi, and R. G. Vickson. “A survey of the maximum principles for optimal control problems with state constraints”. In: *SIAM Review* 37.2 (1995), pp. 181–218 (cit. on pp. 13, 14, 156).

[56] J. Hasenbein and D. Perry (Eds). “Special Issue on Queueing Systems with Abandonments”. In: *Queueing Systems* 75.2–4 (2013), pp. 111–384 (cit. on p. 7).

[57] F. Iravani and B. Balcioğlu. “On priority queues with impatient customers”. In: *Queueing Systems* 58 (2008), pp. 239–260 (cit. on p. 7).

[58] A. Jean-Marie and E. Hyon. “Scheduling services in a queuing system with impatience and setup costs”. In: *Lecture Notes in Electrical Engineering* 62 (2010), pp. 45–50 (cit. on p. 132).

[59] B. Ji, G.R. Gupta, M Sharma, X. Lin, and N.B. Shroff. “Achieving optimal throughput and near-optimal asymptotic delay performance in multichannel wireless networks with low complexity: a practical greedy scheduling policy”. In: *IEEE/ACM Transactions on Networking* 23.3 (2015), pp. 880–893 (cit. on p. 9).

[60] D.G. Kendall. “Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain”. In: *The Annals of Mathematical Statistics* 24 (1953) (cit. on p. 2).

[61] J. Kim and A. Ward. “Dynamic scheduling of a GI/GI/1+GI queue with multiple customer classes”. In: *Queueing Systems* 75.2-4 (2013), pp. 339–384 (cit. on p. 7).

[62] L. Kleinrock. *Queueing Systems, vol. 1*. John Wiley and Sons, 1976 (cit. on p. 2).

[63] L. Kleinrock. *Queueing Systems, vol. 2*. John Wiley and Sons, 1976 (cit. on p. 2).

[64] L. Kruk, J. Lehoczky, K. Ramanan, and S. Shreve. “Heavy traffic analysis for EDF queues with reneging”. In: *The Annals of Applied Probability* 21.2 (2011), pp. 484–545 (cit. on p. 132).

[65] K. Liu, R. Weber, and Q. Zhao. “Indexability and whittle index for restless bandit problems involving reset processes”. In: *IEEE Conference on Decision and Control and European Control Conference* (2011), pp. 7690–7696 (cit. on p. 10).

[66] K. Liu and Q. Zhao. “Indexability of Restless Bandit Problems and Optimality of Whittle Index for Dynamic Multichannel Access”. In: *IEEE Transactions Information Theory* 56 (2010), pp. 5547–5567 (cit. on p. 10).

[67] C. Maglaras. “Discrete-review policies for scheduling stochastic networks: Trajectory tracking and fluid-scale asymptotic optimality”. In: *The Annals of Applied Probability* 10 (2000), pp. 897–929 (cit. on pp. 10, 113).

[68] S.P. Meyn. “Stability and optimization of queueing networks and their fluid models”. In: *Mathematics of stochastic manufacturing systems. Lectures in Applied Mathematics* 33 (1997), pp. 175–199 (cit. on pp. 10, 113).

[69] J. A. van Mieghem. “Dynamic Scheduling with Convex Delay Costs: The Generalized $c\mu$ Rule”. In: *The Annals of Applied Probability* 5.3 (1995), pp. 808–833 (cit. on pp. 6, 44, 76).

[70] J. Niñ-Mora. “Dynamic priority allocation via restless bandit marginal productivity indices”. In: *TOP* 15.2 (2007), pp. 161–198 (cit. on pp. 9, 26).

[71] J. Niño-Mora. “Restless Bandits, partial conservation laws and indexability”. In: *Advances in Applied Probability* 33.1 (2001), pp. 76–98 (cit. on p. 4).

[72] J. Niño-Mora and S. Villar. “Sensor scheduling for hunting elusive hiding targets via whittle’s restless bandit index policy”. In: *5th International Conference on Network Games, Control and Optimization (NetGCooP)* (2011), pp. 1–8 (cit. on p. 9).

[73] W. Ouyang, A. Eryilmaz, and N.B. Shroff. “Asymptotically Optimal Downlink Scheduling over Markovian Fading Channels”. In: *Proceedings of IEEE INFOCOM* (2012), pp. 1–9 (cit. on p. 9).

[74] K.P. Papadaki and W.B. Powell. “Exploiting structure in adaptive dynamic programming algorithms for a stochastic batch service problem”. In: *European Journal of Operational Research* 142 (2002), pp. 108–127 (cit. on p. 132).

[75] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko. *The Mathematical Theory of Optimal Processes*. Interscience Publishers, John Wiley and Sons, Inc. New York, 1962 (cit. on p. 14).

[76] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2005 (cit. on pp. 8, 12, 56, 73, 78, 143, 155).

[77] V. Raghunathan, V. Borkar, M. Cao, and P.R. Kumar. “Index Policies for Real-Time Multicast Scheduling for Wireless Broadcast Systems”. In: *Proceedings of IEEE Infocom* (2008) (cit. on p. 10).

[78] Ph. Robert. *Stochastic Networks and Queues*. Springer-Verlag, 2003 (cit. on p. 10).

[79] S. Ross. *Introduction to stochastic dynamic programming*. Academic Press, 1983 (cit. on pp. 12, 155).

[80] A. Salch. “Ordonnancement stochastique avec impatience”. In: *PhD Thesis* (2013) (cit. on p. 7).

[81] A. Salch, J.P. Gayon, and P. Lemaire. “Optimal control of a multi-class queueing system with abandonments.” In: *MSOM, Fontainebleau* (2013) (cit. on p. 7).

[82] L.E. Schrage and L.W. Miller. “The Queue M/G/1 with the Shortest Remaining Processing Time Discipline”. In: *Operations Research* 14 (1966), pp. 670–684 (cit. on p. 20).

[83] A. Seierstad and K. Sydsæter. “Optimal control theory with economic applications”. In: *Advanced textbooks in economics* (1987) (cit. on p. 156).

[84] R. Singh, X. Guo, and P.R. Kumar. “Index policies for optimal mean-variance trade-off inter-delivery times in real-time sensor networks”. In: *Proceedings of IEEE Infocom* (2015) (cit. on p. 10).

[85] I. Taboada, F. Liberal, and P. Jacko. “An opportunistic and non-anticipating size-aware scheduling proposal for mean holding cost minimization in time-varying channels.” In: *Performance Evaluation* 79 (2014), pp. 90–103 (cit. on p. 9).

[86] I. Taboada, P. Jacko, U. Ayesta, and F. Liberal. “Opportunistic scheduling of flows with general size distribution in wireless time-varying channels”. In: *Proceedings of 26th International Teletraffic Congress* (2014) (cit. on p. 9).

[87] H. Takagi. *Queueing analysis: a foundation of performance evaluation*. Queueing Analysis: A Foundation of Performance Evaluation vol. 2. North-Holland, 1993 (cit. on p. 2).

[88] D. Towsley and S.S. Panwar. “On the optimality of Minimum Laxity and Earliest Deadline First scheduling for real-time multiprocessors”. In: *Proceedings of IEEE EUROMICRO-90 Real Time Workshop* (1990), pp. 17–24 (cit. on p. 132).

[89] E. Trélat. “Contrôle optimal: théorie et applications”. In: *Vuibert, Collection : Mathématiques Concrètes* (2008) (cit. on p. 13).

[90] M.H. Veatch and L.M. Wein. “Scheduling a make-to-stock queue: index policies and hedging points”. In: *Operations Research* 44.4 (1996), pp. 634–647 (cit. on pp. 7, 24, 87).

[91] M.H. Veatch and L.M. Wein. "Optimal dynamic scheduling policy for a make-to-stock production system". In: *Operations Research* 45.1 (1997), pp. 42–53 (cit. on p. 7).

[92] I.M. Verloop. "Asymptotically optimal priority policies for indexable and non-indexable restless bandits". In: *HAL Report, hal-00743781* (2015) (cit. on pp. 9, 45).

[93] I.M. Verloop and R. Núñez-Queija. "Asymptotically optimal parallel resource assignment with interference". In: *Queueing Systems* 65.1 (2010), pp. 43–92 (cit. on pp. 10, 113).

[94] R.R. Weber and S.Jr. Stidham. "Optimal Control of Service Rates in Networks of Queues". In: *Advances in Applied Probability* 19.1 (1987), pp. 202–2018 (cit. on p. 6).

[95] R.R. Weber and G. Weiss. "On an index policy for restless bandits". In: *Journal of Applied Probability* 27 (1990), pp. 637–648 (cit. on pp. 4, 9, 45).

[96] G. Weiss. "On optimal draining of reentrant fluid lines". In: *Stochastic Networks, eds. F.P. Kelly and R.J. Williams* (1995), pp. 91–103 (cit. on pp. 10, 71).

[97] W. Whitt. "Efficiency-Driven Heavy-Traffic Approximations for Many-Server Queues with Abandonments". In: *Management Science* 50 (2004), pp. 1449–1461 (cit. on p. 7).

[98] P. Whittle. "Restless Bandits: Activity Allocation in a Changing World". In: *Journal of Applied Probability* 25 (1988), pp. 287–298 (cit. on pp. 3, 8, 9, 22, 24, 27).

[99] A. Wierman, L.L.H. Andrew, and A. Tang. "Power-aware speed scaling in processor sharing systems". In: *Proceedings of IEEE Infocom* (2009) (cit. on pp. 22, 83).